## SEQUENCE LISTING

<110> Wilson, James M. Xiao, Weidong	
<120> Adeno-Associated Virus Serotype I Nucleic Acid Sequences, Vectors and Host Cells Containing Same	
<130> GNVPN.031USA	
<140> US 09/807,802 <141> 2001-04-17	
<150> US 60/107,114 <151> 1998-11-05	
<150> PCT/US99/25694 <151> 1999-11-02	
<160> 20	
<170> PatentIn version 3.1	
<210> 1 <211> 4718	
<212> DNA	
<213> AAV-1	
<220> <221> CDS <222> (335)(2206) <223>	
<220> <221> CDS <222> (2223)(4430) <223>	
<400> 1 ttgcccactc cctctctgcg cgctcgctcg ctcggtgggg cctgcggacc aaaggtccgc	60
agacggcaga gctctgctct gccggcccca ccgagcgagc gagcgcgcag agagggagtg	120
ggcaactcca tcactagggg taatcgcgaa gcgcctccca cgctgccgcg tcagcgctga	180
cgtaaattac gtcatagggg agtggtcctg tattagctgt cacgtgagtg cttttgcgac	240
attttgcgac accacgtggc catttagggt atatatggcc gagtgagcga gcaggatctc	300
cattttgacc gcgaaatttg aacgagcagc agcc atg ccg ggc ttc tac gag atc Met Pro Gly Phe Tyr Glu Ile 1	355
gtg atc aag gtg ccg agc gac ctg gac gag cac ctg ccg ggc att tct Val Ile Lys Val Pro Ser Asp Leu Asp Glu His Leu Pro Gly Ile Ser 10 15 20	403
gac tcg ttt gtg agc tgg gtg gcc gag aag gaa tgg gag ctg ccc ccg Asp Ser Phe Val Ser Trp Val Ala Glu Lys Glu Trp Glu Leu Pro Pro 25 30 35	451

gat Asp	tct Ser	gac Asp	atg Met	Asp	ctg Leu 45	aat Asn	ctg Leu	att Ile	gag Glu	cag Gln 50	gca Ala	ccc Pro	ctg Leu	acc Thr	gtg Val 55		499
40 gcc Ala	gag Glu	aag Lys	ctg Leu	ca.c	כמכ	gac Asp	ttc Phe	ctg Leu	gtc Val 65	caa Gln	tgg Trp	cgc Arg	cgc Arg	gtg Val 70	agt Ser		547
aag Lys	gcc Ala	ccg Pro	gag Glu 75	gcc Ala	ctc Leu	ttc Phe	ttt Phe	gtt Val 80	cag Gln	ttc Phe	gag Glu	aag Lys	ggc Gly 85	gag Glu	tcc Ser		595
tac Tyr	ttc Phe	cac His 90	ctc Leu	cat His	att Ile	ctg Leu	gtg Val 95	gag Glu	acc Thr	acg Thr	Gly	gtc Val 100	aaa Lys	tcc Ser	atg Met		643
gtg Val	ctg Leu 105	ggc Gly	cgc Arg	ttc Phe	ctg Leu	agt Ser 110	cag Gln	att Ile	agg Arg	gac Asp	aag Lys 115	ctg Leu	gtg Val	cag Gln	acc Thr	<u>-</u>	691
atc Ile 120	tac Tyr	cgc Arg	Gly	atc Ile	gag Glu 125	ccg Pro	acc Thr	ctg Leu	ccc Pro	aac Asn 130	тrр	ttc Phe	gcg Ala	gtg Val	acc Thr 135		739
aag Lys	acg Thr	cgt Arg	aat Asn	ggc Gly 140	gcc Ala	gga Gly	ggg	Gly	aac Asn 145	aag Lys	gtg Val	gtg Val	gac Asp	gag Glu 150	tgc Cys		787
tac Tyr	atc Ile	ccc Pro	aac Asn 155	tac Tyr	ctc Leu	ctg Leu	ccc Pro	aag Lys 160	act Thr	cag Gln	ccc Pro	gag Glu	ctg Leu 165	cag Gln	tgg Trp		835
gcg Ala	tgg Trp	act Thr 170	Asn	atg Met	gag Glu	gag Glu	tat Tyr 175	ata Ile	agc Ser	gcc Ala	tgt Cys	ttg Leu 180	. Asii	ctg Leu	gcc Ala		883
gag Glu	cgc Arg 185	Lys	cgg Arg	ctc Leu	gtg Val	gcg Ala 190	cag Gln	cac His	ctg Leu	acc Thr	cac His 195	VUL	agc Ser	caç Glr	acc Thr		931
cag Gln 200	Glu	cag Gln	aac Asn	aag Lys	gag Glu 205	Asn	ctg Leu	aac Asn	Pro	aat Asr 210	1 261	gac Asp	gcg Ala	cct Pro	gtc Val 215		979
atc Ile	cgg Arg	tca Ser	aaa Lys	acc Thr 220	Ser	gcg Ala	cgc Arg	tac Tyr	atg Met 225	GIL	g cto u Lev	g gto l Val	r el7 e aac	y tgg y Trp 230	g ctg b Leu		1027
gtg Val	gac Asp	cgç Arç	g ggc g Gly 235	/ Ile	acc Thr	tco Ser	gag Glu	aag Lys 240	GII	ı tgo	g ato p Ile	c caq e Gli	g gaq n Glu 245	1 110	c cag o Gln		1075
gcc Ala	tco Sei	tac Ty:	r Ile	c tco e Sei	tto Phe	aac Asr	gco Ala 255	AL6	tco a Sei	aa Ası	c tco n Se	g cge r Are 26	y se.	c cae	g atc n Ile		1123
aag Lys	g gco s Ala 26	a Ala	t cto	g gad u Asp	c aat o Asi	gco n Ala 270	a GT	c aad y Lys	g ato	c at e Me	g gc t Al 27	a Le	g ac	c aa r Ly	a tcc s Ser		1171

																	•	
	: :	Ala 280	Pro	gac Asp	Tyr	Leu	Val 1 285	стА	PIO	ALG	110	290				-	295	1219
		aac Asn	cgc Arg	atc Ile	tac Tyr	cgc Arg 300	atc Ile	ctg Leu	gag Glu	ctg Leu	aac Asn 305	ggc Gly	tac Tyr	gaa Glu	cct Pro	gcc Ala 310	tac Tyr	1267
•		gcc Ala	ggc Gly	tcc Ser	gtc Val 315	ttt Phe	ctc Leu	ggc Gly	tgg Trp	gcc Ala 320	cag Gln	aaa Lys	agg Arg	ttc Phe	ggg Gly 325	aag Lys	cgc Arg	1315
		aac Asn	acc Thr	atc Ile 330	tgg Trp	ctg Leu	ttt Phe	GJA aaa	ccg Pro 335	gcc Ala	acc Thr	acg Thr	ggc Gly	aag Lys 340	acc Thr	aac Asn	atc Ile	1363
		gcg Ala	gaa Glu 345	gcc Ala	atc Ile	gcc Ala	cac His	gcc Ala 350	gtg Val	ccc Pro	ttc Phe	tac Tyr	ggc Gly 355	tgc Cys	gtc Val	aac Asn	tgg Trp	1411
		acc Thr 360	Asn	gag Glu	aac Asn	ttt Phe	ccc Pro 365	ttc Phe	aat Asn	gat Asp	tgc Cys	gtc Val 370	gac Asp	aag Lys	atg Met	gtg Val	atc Ile 375	1459
		tgg Trp	tgg Trp	gag Glu	gag Glu	ggc Gly 380	aag Lys	atg Met	acg Thr	gcc Ala	aag Lys 385	gtc Val	gtg Val	gag Glu	tcc Ser	gcc Ala 390	aag Lys	1507
		gcc Ala	att	ctc Leu	ggc Gly 395	ggc Gly	agc Ser	aag Lys	gtg Val	cgc Arg 400	gtg Val	gac Asp	caa Gln	aag Lys	tgc Cys 405	aag Lys	tcg Ser	1555
		tco Ser	gcc Ala	cag Gln 410	Ile	gac Asp	ccc Pro	acc Thr	ccc Pro 415	val	atc Ile	gtc Val	acc Thr	tcc Ser 420		acc Thr	aac Asn	1603
	٠	atç Met	tgo Cys	c gcc s Ala	gtg Val	att Ile	gac Asp	ggg Gly 430	ASI	agc Ser	acc Thr	acc	tto Phe 435		cac His	cag Gln	cag Gln	1651
	:	cco Pro	Le	u Glr	g gac n Asp	cgg Arg	atg Met 445	Phe	aaa Lys	ttt Phe	gaa Glu	cto Lev 450		cgc Arg	cgt Arg	cto Lev	gag Glu 455	1699
		ca <sup>t</sup> Hi:	t ga s As	c ttt p Phe	ggc Gly	: aag , Lys 460	; Val	aca Thi	a aag c Lys	g cag s Glr	g gaa Glu 465	, va.	aaa L Lys	a gaç s Glu	tto Phe	2 tto 2 Phe 470	cgc Arg	1747
		tg: Tr	g gc p Al	g caq a Gli	g gat n Asp 475	His	gtç Val	g acc	c gaq r Gl	g gto u Val 480	LATO	g cat a His	gaq s Gl	g tto u Phe	tac Ty:	_	aga L Arg	1795
		aa Ly	g gg s Gl	t gg y Gl 49	y Ala	c aac a Ası	c aaa n Lys	a ag	a cc g Pr 49	O AL	c cc a Pr	c ga o Asj	t ga p As	c gco p Ala 50	1	t aa p Ly	a agc s Ser	1843
	٠	ga Gl	g cc u Pr 50	o Ly	g cg s Ar	g gc	c tgo a Cy:	c cc s Pr 51	o se	a gt r Va	c gc	g ga a As	t cc p Pr 51	0 50	g ac	g tc r Se	a gac r Asp	1891

gcg Ala 520	Glu	Gly	Ala	Pro	525	Asp	rne	ΑΙα	ASP	530	-1-				5	35	1939
tct Ser	cgt Arg	cac His	gcg Ala	ggc Gly 540	atg Met	ctt Leu	cag Gln	atg Met	ctg Leu 545	ttt Phe	ccc Pro	tgc Cys	aag Lys	aca Thr 550	t C	ys Ys	1987
gag Glu	aga Arg	atg Met	aat Asn 555	cag Gln	aat Asn	ttc Phe	aac Asn	att Ile 560	tgc Cys	ttc Phe	acg Thr	cac His	ggg Gly 565	acg Thr	a A	ga irg	2035
gac Asp	tgt Cys	tca Ser 570	Glu	tgc Cys	ttc Phe	ccc Pro	ggc Gly 575	gtg Val	tca Ser	gaa Glu	tct Ser	caa Gln 580		gto Val	e g L V	ntc Val	2083
aga Arg	aag Lys 585	Arg	acg Thr	tat Tyr	cgg Arg	aaa Lys 590	ctc Leu	tgt Cys	gcc Ala	att Ile	cat His 595		ctg Leu	Le	a (	ely gag	2131
cgg Arg 600	gct Ala		gag Glu	ı att	gct Ala 605	tgc Cys	tcg Ser	gcc Ala	tgc Cys	gat Asp 610		g gtc Val	aac Asn	gto Va	g (	gac Asp 615	2179
cta	aat	gac Asp	tgt Cys	gtt Val	Ser	gag Glu	caa Gln	taa	atg	actt	aaa	ccaç	ggt a N		gc Al 62	t gcc a Ala 5	2231
gat Asp	ggt Gl	tat Ty	ctt Leu 630	cca ı Pro	gat Asp	tgg Trp	ctc Leu	gag Glu 635	ASP	aac Asn	: cto Lev	c tct ı Sei	gaq Glu 640		c Y	att Ile	2279
cgc Arg	gaq Glu	g tgo ı Tr	o Tr	g gad p Asp	ttg Lev	ı aaa Lys	cct Pro	у Сту	gcc Ala	ccg Pro	g aaq b Ly:	g ccc s Pro 65		a go s Al	c a	aac Asn	2327
caç Glr	g caa n Gl: 66	n Ly	g ca s Gl	g gao n Asj	c gad o Asp	665 665	Ari	g ggt g Gly	cto Lev	g gto 1 Va	g ct l Le 67		t gg o Gl	c ta y Ty	r	aag Lys	2375
ТУ	r Le	u Gl	y Pr	c tt o Ph	e Ası	u GT	a cto y Le	c gad u Asp	aaq b Lys	g gg s Gl; 68	y	g cc u Pr	c gt o Va	c aa l As	ac sn	gcg Ala 690	2423
				g gc a Al 69	c ct a Le	c ga u Gl	g ca u Hi	c ga s As	c aa p Ly 70	3 A.	c ta a Ty	c ga r As	c ca p Gl	g ca n G.	ag ln 05	ctc Leu	2471
aa Ly	a go s Al	g gg a Gl	gt ga .y As 71	ıc aa sp As	t cc n Pr	g ta o Ty	c ct r Le	g cg u Ar 71	9 - 1	t aa r As	c ca n Hi	ic go .s Al	cc ga .a As 72	ic g sp A 20	cc la	gag Glu	2519
tt Ph	t ca e Gl	n G.	ag co Lu Ai 25	gt ct cg Le	g ca eu Gl	a ga n Gl	a ga u As 73	b m	g to r Se	t tt r Ph	t gg ie Gl	- 3	gc aa Ly As 35	ac c sn L	tc eu	G] A aaa	2567
cg Ar	g A.	ca g La V	tc to	tc ca he Gl	ag go ln Al	c aa .a Ly .74	ւջ Իշ	ag cg ys Ar	g gt g Va	t ct	-u	aa co lu P: 50	ct c ro L	tc g eu G	igt Sly	ctg Leu	2615

gtt g Val G 755	ag g lu G	jaa ( ilu (	ggc ( Gly /	Ala .	aag Lys 760	acg Thr	gct Ala	cct Pro	gga Gly	aag Lys 765	aaa Lys	cgt Arg	ccg Pro	gta Val	gag Glu 770		2663
cag t Gln S	cg o	cca Pro	Gln	gag Glu 775	cca Pro	gac Asp	tcc Ser	tcc Ser	tcg Ser 780	ggc Gly	atc Ile	ggc Gly	aag Lys	aca Thr 785	ggc Gly		2711
cag c Gln G	ag (	Pro	gct Ala 790	aaa Lys	aag Lys	aga Arg	ctc Leu	aat Asn 795	ttt Phe	ggt Gly	cag Gln	act Thr	ggc Gly 800	gac Asp	tca Ser		2759
gag t Glu S	Ger '	gtc Val 805	ccc Pro	gat Asp	cca Pro	caa Gln	cct Pro 810	ctc Leu	gga Gly	gaa Glu	cct Pro	cca Pro 815	gca Ala	acc Thr	ccc Pro		2807
gct ç Ala <i>A</i>	gct Ala 320	gtg Val	gga Gly	cct Pro	act Thr	aca Thr 825	atg Met	gct Ala	tca Ser	ggc Gly	ggt Gly 830	4	gca Ala	cca Pro	atg Met		2855
gca 9 Ala <i>P</i> 835	gac Asp	aat Asn	aac Asn	gaa Glu	ggc Gly 840	gcc Ala	gac Asp	gga Gly	gtg Val	ggt Gly 845		gcc Ala	tca Ser	gga Gly	aat Asn 850		2903
tgg ( Trp l	cat His	tgc Cys	gat Asp	tcc Ser 855	aca Thr	tgg Trp	ctg Leu	ggc Gly	gac Asp 860	1119	gto Val	ato Ile	acc Thr	acc Thr 865	agc Ser		2951
acc o	cgc Arg	acc Thr	tgg Trp 870	gcc Ala	ttg Leu	ccc Pro	acc	tac Tyr 875	L'311	. aac Asn	cac His	cto Lev	tac Tyr 880	aag Lys	g caa Gln		2999
atc Ile	tcc Ser	agt Ser 885	gct Ala	tca Ser	acg Thr	gly	gcc Ala 890	, ber	aac Asr	gac Asp	aac Ası	c cac n His 895	tac Tyi	tto Phe	e ggc		3047
Tyr	agc Ser 900	acc Thr	ccc Pro	tgg Trp	ggg Gly	tat Tyi	PHE	gat Asp	tto Phe	aac Ası	aga n Are 91	J	c cad e Hi:	c tg s Cy	c cac s His		3095
ttt Phe 915	tca Ser	cca Pro	. cgt Arg	gac Asp	tgo Trp 920	) GTI	g cga n Arg	a cto	ato u Ilo	c aac e Asi 92		c aa n As	t tg n Tr	b ej a aa	a ttc y Phe 930		3143
cgg Arg	ccc Pro	aag Lys	g aga Arg	tet J Let 935	ı Ası	tte n Ph	c aaa e Ly	a cto s Le	c tt u Ph 94		c at n Il	c ca e Gl	a gt n Va	с аа 1 Ly 94	g gag s Glu 5		3191
gtc Val	acg Thr	acç Thi	g aat c Asi 950	n Asl	t gg p Gl	c gt y Va	c ac l Th	a ac r Th 95		c gc e Al	t aa a As	it aa sn As	c ct n Le 96	t ac u Th	c ago ir Sei	3 °.	3239
acg Thr	gtt Val	caa Gli 96	n Va	c tt l Ph	c tc e Se	g ga r As	c tc p Se 97	L GI	g ta u Ty	c ca r Gl	g ct n Le	eu Pr 97	_	ıc gt vr Va	c cto	: 1	3287
G1 y	tct Sei 980	: Al	g ca a Hi	c ca s Gl	g gg n Gl	c tg y Cy 98	'S Le	c cc eu Pr	t co	g tt		eg go ro Al 90	eg ga La As	ac gi	tg tto al Pho	c e	3335

•	atg att ccg caa tac ggc tac ctg acg ctc aac aat ggc agc caa Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly Ser Gln 995 1000	3380
	gcc gtg gga cgt tca tcc ttt tac tgc ctg gaa tat ttc cct tct Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro Ser 1010 1015	3425
	cag atg ctg aga acg ggc aac aac ttt acc ttc agc tac acc ttt Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe 1025	3470
	gag gaa gtg cct ttc cac agc agc tac gcg cac agc cag agc ctg Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu 1040 1045	3515
	gac cgg ctg atg aat cct ctc atc gac caa tac ctg tat tac ctg Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu 1055 1060 1065	3560
	aac aga act caa aat cag too gga agt goo caa aac aag gac ttg Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu 1070 1075	3605
	ctg ttt agc cgt ggg tct cca gct ggc atg tct gtt cag ccc aaa Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys 1085	3650
	aac tgg cta cct gga ccc tgt tat cgg cag cag cgc gtt tct aaa Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys 1100 1105	3695
	aca aaa aca gac aac aac agc aat ttt acc tgg act ggt gct Thr Lys Thr Asp Asn Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala 1115 1120 1125	3740
	tca aaa tat aac ctc aat ggg cgt gaa tcc atc atc aac cct ggc Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly 1130 1135	378
	act gct atg gcc tca cac aaa gac gac gaa gac aag ttc ttt ccc Thr Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro 1145 1150 1155	383
	atg agc ggt gtc atg att ttt gga aaa gag agc gcc gga gct tca Met Ser Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser 1160 1165	387
	aac act gca ttg gac aat gtc atg att aca gac gaa gag gaa att Asn Thr Ala Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile 1175 1180 1185	392
	aaa gcc act aac cct gtg gcc acc gaa aga ttt ggg acc gtg gca Lys Ala Thr Asn Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala 1190 1195	396
	gtc aat ttc cag agc agc aca gac cct gcg acc gga gat gtg Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala Thr Gly Asp Val 1205 1210 1215	401

cat gct atg gga gca tta cct ggc atg gtg tgg caa gat aga gac His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln Asp Arg Asp 1220	4055
gtg tac ctg cag ggt ccc att tgg gcc aaa att cct cac aca gat Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His Thr Asp 1245	4100
gga cac ttt cac ccg tct cct ctt atg ggc ggc ttt gga ctc aag Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu Lys 1250	4145
aac ccg cct cct cag atc ctc atc aaa aac acg cct gtt cct gcg Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala	4190
aat cct ccg gcg gag ttt tca gct aca aag ttt gct tca ttc atc  Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile  1285	4235
acc caa tac tcc aca gga caa gtg agt gtg gaa att gaa tgg gag Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu 1305	4280
ctg cag aaa gaa aac agc aag cgc tgg aat ccc gaa gtg cag tac Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr	4325
aca tcc aat tat gca aaa tct gcc aac gtt gat ttt act gtg gac Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp	4370
aac aat gga ctt tat act gag cct cgc ccc att ggc acc cgt tac Asn Asn Gly Leu Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr	4415
ctt acc cgt ccc ctg taattacgtg ttaatcaata aaccggttga ttcgtttc Leu Thr Arg Pro Leu	ag 4470
1355 ttgaactttg gtctcctgtc cttcttatct tatcggttac catggttata gcttaca	acat 4530
taactgcttg gttgcgcttc gcgataaaag acttacgtca tcgggttacc cctagt	gatg 4590
gagttgccca ctccctctct gcgcgctcgc tcgctcggtg gggcctgcgg accaaa	ggtc 4650
gagttgccca ctccctctct gegegeeess 5	ggga 4710
	4718
gtgggcaa	

<210> 2 <211> 623 <212> PRT <213> AAV-1

<400> 2
Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp
10
15

- Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu 20 25 30
- Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile 35 40 45
- Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu 50 60
- Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val 65 70 75 80
- Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Ile Leu Val Glu 85 90
- Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile 100 105 110
- Arg Asp Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu 115 120 125
- Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly 130 140
- Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys 145 150 155 160
- Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile 165 170 175
- Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His 180 185
- Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn 195 200 205
- Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr 210 220
- Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys 235 240
- Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala 255 255
- Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys 260 265 270
- Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ala 275 280 285
- Pro Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu 290 295 300
- Asn Gly Tyr Glu Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala 305 310 315
- Gln Lys Arg Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala 325 330 335

Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro 340 345 350

Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp 365

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 370 380

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 385 390 400

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val 405 410 415

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 420 425 430

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe 435

Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln
450 455

Glu Val Lys Glu Phe Phe Arg Trp Ala Gln Asp His Val Thr Glu Val 465 470 475

Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Asn Lys Arg Pro Ala 485 490 495

Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val 500 505 510

Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala 515 520 525

Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Leu Gln Met 530 540

Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile 545 550 550

Cys Phe Thr His Gly Thr Arg Asp Cys Ser Glu Cys Phe Pro Gly Val 565 570 575

Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys 580 585

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala 595 600 605

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln 610 620

<210> 3

<211> 736

<212> PRT

<213> AAV-1

- Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro 20 25 30
- Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro 35 40 45
- Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro 50 60
- Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp 65 70 75 80
- Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala 85 90 95
- Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly 100 105
- Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro 115 120 125
- Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg 130 135
- Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly 145 150 155
- Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr 165 170 175
- Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro 180 185 190
- Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly 195 200 205
- Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala
- Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile 225 230 235
- Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu
  245 250 255
- Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His 260 265 270
- Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe 275 280 285
- His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn 290 295 300
- Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn Ile Gln

Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn Asn 335

- Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu Pro 340 345
- Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala 355 360 365
- Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly 370 380
- Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro 385 390 395
- Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr Phe 405 410 415
- Glu Glu Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp 420 425 430
- Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Asn Arg 445
- Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn Lys Asp Leu Leu Phe Ser 450 455
- Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro 480
- Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn 485 490 495
- Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn 500 505
- Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys 515 520 525
- Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly 530 540
- Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile 545 550 560
- Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg 575
- Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala 580 585 590
- Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp Gln 595 600 605
- Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His 610 620
- Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu

640. 635 630 625 Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala 645 Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile Thr 665 Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Ser Asn Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gly Leu 710 Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pro Leu 730 <210> 1872 <211> DNA <212> <213> AAV-1 <220> CDS <221> (1)..(1869)<222> <223> atg ccg ggc ttc tac gag atc gtg atc aag gtg ccg agc gac ctg gac 48 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp gag cac ctg ccg ggc att tct gac tcg ttt gtg agc tgg gtg gcc gag 96 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu 20 aag gaa tgg gag ctg ccc ccg gat tct gac atg gat ctg aat ctg att 144 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile 40 gag cag gca ccc ctg acc gtg gcc gag aag ctg cag cgc gac ttc ctg 192 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu gtc caa tgg cgc cgc gtg agt aag gcc ccg gag gcc ctc ttc ttt gtt 240 Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val cag ttc gag aag ggc gag tcc tac ttc cac ctc cat att ctg gtg gag 288

Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Ile Leu Val Glu

acc acg ggg gtc aaa tcc atg gtg ctg ggc cgc ttc ctg agt cag att

Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile 105

100

agg Arg	gac Asp	aag Lys 115	ctg Leu	gtg Val	cag Gln	Thr	atc Ile 120	tac Tyr	cgc Arg	eja gaa	atc Ile	gag Glu 125	ccg Pro	acc Thr	: C	tg eu	384
Pro	Asn 130	Trp	Phe	Ala	vai	acc Thr 135	гур	1111	ALG	11011	140		-	_			432
aac Asn 145	aag Lys	gtg Val	gtg Val	gac Asp	gag Glu 150	tgc Cys	tac Tyr	atc Ile	ccc Pro	aac Asn 155	tac Tyr	ctc Leu	ctg Leu	Pro	e a o I	aag Lys 160	480
act Thr	cag Gln	ccc Pro	gag Glu	ctg Leu 165	cag Gln	tgg Trp	gcg Ala	tgg Trp	act Thr 170	aac Asn	atg Met	gag Glu	gag Glu	tai Ty: 17:	t a r :	ata Ile	528
agc Ser	gcc Ala	tgt Cys	ttg Leu 180	aac Asn	ctg Leu	gcc Ala	gag Glu	cgc Arg 185	aaa Lys	cgg Arg	ctc Leu	gtg Val	gcg Ala 190		g o	cac His	576
ctg Leu	acc Thr	cac His 195	Val	agc Ser	cag Gln	acc Thr	cag Gln 200	gag Glu	cag Gln	aac Asn	aag Lys	gag Glu 205		ct Le	g u	aac Asn	624
ccc Pro	aat Asn 210	tct Ser	gac Asp	gcg Ala	cct Pro	gtc Val 215	atc Ile	cgg Arg	tca Ser	aaa Lys	acc Thr 220		gcg Ala	cg Ar	g	tac Tyr	672
atg Met 225	Glu	ctg Leu	gtc Val	Gly	tgg Trp 230	ctg Leu	gtg Val	gac Asp	cgg Arg	ggc Gly 235		acc Thr	: tcc :Ser	ga Gl	g .u	aag Lys 240	720
cag Gln	tgg Trp	ato Ile	cag Gln	gag Glu 245	ı Asp	cag Gln	gcc Ala	tcg Ser	tac Tyr 250	110	tco Sei	tto Phe	aac Asr	g gc n Al 25	.c .a .5	gct Ala	768
tcc Ser	aac Asn	tcg Sei	g cgg Arg 260	, Ser	cag Gln	atc Ile	aag Lys	g gcc Ala 265	MIG	ctg Leu	gaq 1 Asj	c aat o Asi	gce n Ala 27		gc Ly	aag Lys	816
ato Ile	atç Met	gcg Ala 27	a Lei	g aco ı Thi	c aaa c Lys	tcc Ser	gcg Ala 280	, FIC	gac Asp	tac Tyi	c cto	g gta u Va 28		y Pi	ro	gct Ala	864
ccq Pro	g cco Pro 290	Al	g gad a Asj	c att	t aaa e Lys	a acc Thr 295	AS	c cgo n Aro	ato g Ile	tac Ty:	c cg r Ar 30	9	c ct e Le	g ga u G	ag lu	ctg Leu	912
aa As: 30:	n Gl	c ta y Ty	c ga r Gl	a cc u Pr	t gco o Ala 31	tac a Tyi	gc Al	c gg a Gl	c tc y Se	c gt r Va 31	<b>-</b>	t ct e Le	c gg u Gl	c t y T	gg rp	gcc Ala 320	960
ca Gl:	g aa n Ly	a ag s Ar	g tt g Ph	c gg e Gl 32	А ГА	g cgo s Aro	c aa g As	c ac n Th	c at r Il 33	C 11	g ct p Le	g tt eu Ph	t gg e Gl	- 2	cg ro 35	gcc Ala	1008
ac Th	c ac r Th	g gg r Gl	с аа у Ly 34	's Th	c aa r As	c at n Il	c gc e Al	g ga a Gl 34	u Ai	c at a Il	c go e Al	cc ca La Hi		cc g La V 50	jtg /al	ccc Pro	1056

						•												
	ttc Phe	tac Tyr	ggc Gly 355	tgc Cys	gtc Val	aac Asn	tgg Trp	acc Thr 360	aat Asn	gag Glu	aac Asn	ttt Phe	ccc Pro 365	ttc Phe	aat Asn	gat Asp		1104
	tgc Cys	gtc Val 370	gac Asp	aag Lys	atg Met	gtg Val	atc Ile 375	tgg Trp	tgg Trp	gag Glu	gag Glu	ggc Gly 380	aag Lys	atg Met	acg Thr	gcc Ala		1152
	aag Lys 385	gtc Val	gtg Val	gag Glu	tcc Ser	gcc Ala 390	aag Lys	gcc Ala	att Ile	ctc Leu	ggc Gly 395	Gly	agc Ser	aag Lys	gtg Val	cgc Arg 400		1200
	Val	Asp	Gln	Lys	Cys 405	Lys	ser	tcc Ser	Ala	410	110	ASP			415		÷	1248
	Ile	Val	Thr	Ser 420	Asn	Thr	Asn	atg Met	425	Ala	vai		ASP	430				1296
	Thr	Thr	Phe 435	Glu	His	GIn	GIN	440	Leu	GIII	vob	nig	445		-4			1344
	gaa Glu	ctc Leu 450	Thr	cgc Arg	cgt Arg	ctg Leu	gag Glu 455	cat His	gac Asp	ttt Phe	ggc	: aag ' Lys 460		aca Thr	aag Lys	cag Gln		1392
	gaa Glu 465	Val	aaa Lys	gag Glu	ttc Phe	ttc Phe 470	Arg	tgg Trp	gcg Ala	cag Gln	gat Asp 475	, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	gtg Val	Thr	gag Glu	gtg Val 480		1440
	gcg Ala	cat His	gag Glu	ttc Phe	tac Tyr 485	Val	aga Arg	aag Lys	ggt Gly	gga Gly 490	AT	aac Ası	aaa Lys	aga Arg	Pro 495	gcc Ala		1488
	ccc Pro	gat Asp	gac Asp	gcg Ala 500	Asp	aaa Lys	agc Ser	gag Glu	ccc Pro 505	, гу	g cgg s Arg	g gc	c tgo a Cys	510		a gtc r Val		1536
	gcg	gat Asp	cca Pro	Ser	acg Thr	tca Ser	gac : Asp	gcg Ala 520	GIL	ı gga	a gct y Ala	t cc a Pr	g gto o Val 52		c tti p Pho	t gcc e Ala		1584
·	gac Asp	agg Arg 530	д Туі	caa Glr	aac Asr	: aaa Lys	tgt Cys 535	s Ser	cgt Arc	cao g Hi	c gc	g gg a Gl 54	y IIC	g ct t Le	t ca u Gl	g atg n Met		1632
	cto Lev 54!	a Phe	t cco	c tgo cys	aaq Lys	g aca Thi 550	r Cy:	c gaç s Glu	g aga	a ate	g aa t As 55	11 61	g aa n As	t tt n Ph	c aa e As	c att n Ile 560		1680
	tg:	c tto	c ace	g cad	= ggg = Gly 56	y Thi	g ag r Ar	a gad g Asj	c tg	t tc s Se 57	T GT	g tg u Cy	c tt s Ph	c cc e Pr	c gg c Gl 57	c gtg y Val 5	J -	1728
	tc Se	a ga r Gl	a tc u Se	t cas r Gl: 58	n Pr	g gt o Va	c gt l Va	c aga l Ara	a aa g Ly 58	S AL	g ac	g ta	it cg /r Ar	g aa g Ly 59	, 5 20	c tgt u Cys	5	1776

230

1824

- Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala 255 255
- Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys 260 265 270
- Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ala 275 280 285
- Pro Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu 290 295
- Asn Gly Tyr Glu Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala 305 310 315
- Gln Lys Arg Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala 325
- Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro 340 345 350
- Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp 365
- Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 370 380
- Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 385 390 395
- Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val 405 410 415
- Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 420 425 430
- Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe
  435 440 445
- Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln 450 455
- Glu Val Lys Glu Phe Phe Arg Trp Ala Gln Asp His Val Thr Glu Val 475 470 475
- Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Asn Lys Arg Pro Ala 495 495
- Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val 500 505 510
- Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala 515
- Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Leu Gln Met 530 540
- Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile 545 550 550 560

Cys Phe Thr His Gly Thr Arg Asp Cys Ser Glu Cys Phe Pro Gly Val 565 Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys 580 Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala 595 Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln 615 610 <210> 6 <211> 1641 DNA <212> <213> AAV-1 <220> CDS <221> (1)..(1638) <222> <223> atg ccg ggc ttc tac gag atc gtg atc aag gtg ccg agc gac ctg gac 48 Met Pro Gly Phe Tyr Glu Ile Val Ile Lys Val Pro Ser Asp Leu Asp gag cac ctg ccg ggc att tct gac tcg ttt gtg agc tgg gtg gcc gag 96 Glu His Leu Pro Gly Ile Ser Asp Ser Phe Val Ser Trp Val Ala Glu aag gaa tgg gag ctg ccc ccg gat tct gac atg gat ctg aat ctg att 144 Lys Glu Trp Glu Leu Pro Pro Asp Ser Asp Met Asp Leu Asn Leu Ile 35 gag cag gca ccc ctg acc gtg gcc gag aag ctg cag cgc gac ttc ctg 192 Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu 50 gtc caa tgg cgc cgc gtg agt aag gcc ccg gag gcc ctc ttc ttt gtt 240 Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val 70 65 cag ttc gag aag ggc gag tcc tac ttc cac ctc cat att ctg gtg gag 288 Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Ile Leu Val Glu acc acg ggg gtc aaa tcc atg gtg ctg ggc cgc ttc ctg agt cag att Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile 336 105 100 agg gac aag ctg gtg cag acc atc tac cgc ggg atc gag ccg acc ctg 384 Arg Asp Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu 120 115 ccc aac tgg ttc gcg gtg acc aag acg cgt aat ggc gcc gga ggg 432 Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly 140 135 130

aac a Asn 1	aag Lys	gtg Val	gtg Val	gac Asp	gag Glu 150	tgc ' Cys '	tac Tyr	atc Ile		aac Asn ' 155	tac Tyr	ctc ( Leu :	ctg ( Leu l	ccc Pro	aag Lys 160		480
act Thr	cag Gln	ccc Pro	gag Glu	ctg Leu 165	cag Gln	tgg Trp	gcg Ala	tgg Trp	act Thr 170	aac Asn	atg Met	gag Glu	gag Glu	tat Tyr 175	ata Ile		528
agc Ser	gcc Ala	tgt Cys	ttg Leu 180	Asn	ctg Leu	gcc Ala	gag Glu	cgc Arg 185	aaa Lys	cgg Arg	ctc Leu	gtg Val	gcg Ala 190	cag Gln	cac His		576
ctg Leu	acc Thr	cac His 195	Val	agc Ser	cag Gln	acc Thr	cag Gln 200	gag Glu	cag Gln	aac Asn	aag Lys	gag Glu 205	aat Asn	ctg Leu	aac Asn		624
ccc Pro	aat Asn 210	Ser	gac Asp	gcg Ala	cct Pro	gtc Val 215	atc Ile	cgg Arg	tca Ser	aaa Lys	acc Thr 220	tcc Ser	gcg Ala	cgc Arg	tac Tyr		672
atg Met 225	gag Glu	ctç Lev	g gto l Val	r eja gaa	tgg Trp 230	ctg Leu	gtg Val	gac Asp	cgg Arg	ggc Gly 235	atc Ile	acc Thr	tcc Ser	gag Glu	aag Lys 240		720
	tgg Trp	ato Ile	caq e Gli	g gag n Glu 245	gac Asp	cag Gln	gcc Ala	tcg Ser	tac Tyr 250		tcc Ser	ttc Phe	aac Asn	gcc Ala 255	gct Ala		768
tcc Ser	aac Asn	tc Se	g cg r Ar 26	g Sei	cag Gln	atc Ile	aag Lys	gcc Ala 265	rinu	ctg Leu	gac Asp	aat Asn	gcc Ala 270	Gl y	aag Lys		816
atc Ile	ato Met	g gcc : Ala 27	a Le	g aco u Th:	c aaa c Lys	tcc Ser	gcg Ala 280	ETC	gac Asp	tac Tyr	ctç Lev	g gta u Val 285	_	Pro	gct Ala	,	864
ccg Pro	CC0 Pro	o Al	g ga a As	c at p Il	t aaa e Lys	acc Thr 295	ASI	c cgo	ato	c tac e Tyr	2 cg( Ar(	-	ctg Leu	gaq Gli	g ctg u Leu		912
aac Asr 305	gg Gl		c ga r Gl	a cc u Pr	t gco o Ala 310	TAI	gc Al	c gg a Gl	c tc y Se.	c gto r Val	_	t cto e Lev	ggc Gly	tg Tr	g gcc p Ala 320	,	960
		a ag s Ar	ıg tt :g Pl	c gg ne Gl 32	у гу	g cgo	c aa g As	c ac n Th	c at r Il 33	,	g ct p Le	g tti u Pho	t ggg e Gl	g cc y Pr 33	g gcc o Ala 5		1008
acc Th	c ac r Th	g gg	LY L	ag ac ys Th 40	c aa r As	c at n Il	c gc e Al	g ga a G1 34	u 111	c at a Il	c gc e Al	c ca .a Hi	c gc s Al 35	c gt a Va 0	g ccc l Pro		1056
tt Ph	c ta e Ty	r G	gc t ly C 55	gc gt ys Va	c aa al As	c tg n Tr	g ac p Th 36	11 /	it ga in Gl	ig aa .u As	c tt n Ph	t cc ne Pr 36	c tt o Ph	c aa e As	at gat sn Asp		1104
tg Cy	's Va	ic g al A 70	ac a sp L	ag a ys M	tg gt et Va	g at 11 I1 37		gg to	gg ga rp Gl	ag ga Lu Gl		gc aa ly Ly 80	ig at /s Me	g ac	og god nr Ala	: 1	1152

aag g Lys V	al '	Val	GLu	Ser	390	цуз .	, Lu			395					400	)	1200
gtg g Val A	jac Asp	caa Gln	aag Lys	tgc Cys 405	aag Lys	tcg Ser	tcc Ser	gcc Ala	cag Gln 410	atc Ile	gac Asp	ccc Pro	acc (	ccc Pro 415	gto Vai	J a	1248
atc g Ile V	gtc Val	acc Thr	tcc Ser 420	aac Asn	acc Thr	aac Asn	atg Met	tgc Cys 425	gcc Ala	gtg Val	att Ile	gac Asp	ggg Gly 430	aac Asn	ag Se	c r	1296
acc a	acc Thr	ttc Phe 435	Glu	cac His	cag Gln	cag Gln	ccg Pro 440	ttg Leu	cag Gln	gac Asp	cgg Arg	atg Met 445	ttc Phe	aaa Lys	tt Ph	t ie	1344
gaa Glu l	ctc Leu 450			cgt Arg	ctg Leu	gag Glu 455	111.5	gac Asp	ttt Phe	ggc Gly	aag Lys 460	gtg Val	aca Thr	aag Lys	ca Gl	ig Ln	1392
		aaa Lys	gag Glu	tto Phe	tto Phe	: ALG	tgg Trp	gcg Ala	cag Gln	gat Asp 475	cac His	gtg Val	acc Thr	gag Glu	gt Va 41	tg al 80	1440
	cat His	gaç Glu	g tto i Phe	tac Tyl	c val	aga Arg	aag Lys	ggt Gl	gga Gly 490		: aac Asr	aaa Lys	aga Arg	Pro 495	g A	cc la	1488
ccc Pro	gat Asp	ga Asj	c gcg o Ala 500	g gat a Asp		a ago s Ser	gaç Glu	g cco 1 Pro 50!		g cgg s Arg	g gco	c tgc a Cys	ccc Pro 510	t ca Se i	ag cV	tc al	1536
gcg Ala	gat Asp	cc Pr	a tc		g tca r Se	a gad r Asp	g gcg Ala 52	<u> </u>	a gg u Gl	a gct y Ala	t cce	g gtg o Val 525	gac Asp	tt! Ph	t g e A	cc la	1584
gac Asp	agg Arg 530	j ta j Ty		c tg y Cy	c cg s Ar	a tg g Tr 53	р пе	a tc u Se	t tc r Se	c ag r Ar	a tt g Le 54	g gct u Ala 0	cga Arg	gg Gl	а с У <sup>С</sup>	aa Hn	1632
cct Pro 545	cto Le	c tg	ıa														1641
<21 <21 <21 <21	l1> l2>	7 546 PRI AAV	r														
1				5					_	_		ro Se					
Glı	u Hi	s L	eu P 2	ro G O	ly I	le S	er A	sp S 2	er P 5	he V	al S	er Tı	cp Va 30	l A	la	Glu	
ГÀ	s G	Lu T 3	rp G	lu L	eu P	ro P	ro A 4	sp S 0	er A	sp M	et A	sp Le	eu As 5	sn L	eu	Ile	

- Glu Gln Ala Pro Leu Thr Val Ala Glu Lys Leu Gln Arg Asp Phe Leu 50 60
- Val Gln Trp Arg Arg Val Ser Lys Ala Pro Glu Ala Leu Phe Phe Val 65 70 75 80
- Gln Phe Glu Lys Gly Glu Ser Tyr Phe His Leu His Ile Leu Val Glu 85 90 95
- Thr Thr Gly Val Lys Ser Met Val Leu Gly Arg Phe Leu Ser Gln Ile 100 105 110
- Arg Asp Lys Leu Val Gln Thr Ile Tyr Arg Gly Ile Glu Pro Thr Leu 125
- Pro Asn Trp Phe Ala Val Thr Lys Thr Arg Asn Gly Ala Gly Gly Gly 130
- Asn Lys Val Val Asp Glu Cys Tyr Ile Pro Asn Tyr Leu Leu Pro Lys 145 150 150
- Thr Gln Pro Glu Leu Gln Trp Ala Trp Thr Asn Met Glu Glu Tyr Ile 165 170 175
- Ser Ala Cys Leu Asn Leu Ala Glu Arg Lys Arg Leu Val Ala Gln His 180 185 190
- Leu Thr His Val Ser Gln Thr Gln Glu Gln Asn Lys Glu Asn Leu Asn 195 200 205
- Pro Asn Ser Asp Ala Pro Val Ile Arg Ser Lys Thr Ser Ala Arg Tyr 210 215 220
- Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys 225 230 235 240
- Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala 245 250 250
- Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys . 260 265 270
- Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ala 275 280 285
- Pro Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu 290 295 300
- Asn Gly Tyr Glu Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala 305 310 315
- Gln Lys Arg Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala 325
- Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro 340 345
- Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp 365

	Cys	Val 370	Asp	Lys	Met	Val	Ile 375	Trp	Trp	Glu	Glu	Gly 380	Lys	Met	Thr	Ala	
	Lys 385	Val	Val	Glu	Ser	Ala 390	Lys	Ala	Ile	Leu	Gly 395	Gly	Ser	Lys	Val	Arg 400	
	Val	Asp	Gln	Lys	Cys 405	Lys	Ser	Ser	Ala	Gln 410	Ile	Asp	Pro	Thr	Pro 415	Val	
٠	Ile	Val	Thr	Ser 420	Asn	Thr	Asn	Met	Cys 425	Ala	Val	Ile	Asp	Gly 430	Asn	Ser	
	Thr	Thr	Phe 435	Glu	His	Gln	Gln	Pro 440	Leu	Gln	Asp	Arg	Met 445	Phe	Lys	Phe	
	Glu	Leu 450		Arg	Arg	Leu	Glu 455	His	Asp	Phe	Gly	Lys 460	Val	Thr	Lys	Gln	
	Glu 465		Lys	Glu	Phe	Phe 470	Arg	Trp	Ala	Gln	Asp 475	His	Val	Thr	Glu	Val 480	
	Ala	His	Glu	Phe	Tyr 485	Val	Arg	Lys	Gly	Gly 490	Ala	Asn	Lys	Arg	Pro 495	Ala	
,	Pro	Asp	Asp	Ala 500	Asp	Lys	Ser	Glu	Pro 505	Lys	Arg	, Ala	Cys	Pro 510	Ser	Val	
	Ala	Asp	Pro 515	Ser	Thr	Ser	Asp	Ala 520	Glu	Gly	Ala	a Pro	Val 525	Asp	Ph∈	Ala	
	Asp	Arg 530		: Gly	Cys	Arg	Trp 535	Leu	Ser	Ser	Arg	540	ı Ala	Arg	ß Gl	/ Gln	
	Pro 545	Leu	1														
			8 1200 DNA AAV-						2.1 2.								
	<22 <22	20> 21> 22> 23>	CDS (1)	(1	197)												
		400> g ga t Gl		g gt u Va	c gg 1 Gl 5	g tg y Tr	g ct p Le	g gt u Va	g ga 1 As	c cg p Ar 10	g Gi	c at y Il	c ac e Th	c tc r Se	c ga r Gl 15	g aag u Lys	48
	ca Gl	g tg n Tr	g at p Il	c ca e Gl 20	n Gl	g ga u As	с са р Gl	g gc n Al	c tc a Se 25	ттА	c at r Il	c to e Se	c tt r Ph	c aa e As 30		c gct a Ala	96
	tc Se	c aa r As	c tc n Se 35	r Ar	g to g Se	c ca r Gl	g at n Il	c aa e Ly 40	SAI	c go a Al	t ct a Le	ig ga eu As	ic aa sp As 45		c gg a Gl	gc aag .y Lys	144

									_		a+ a	at a	ממכ	ccc	act		192
Ile	Met 50	Ala	Leu	Thr		ser 55	Ald	PLO	Asp	ı yı	60		3				
ccg Pro 65	ccc Pro	gcg Ala	gac Asp	att Ile	aaa Lys 70	acc Thr	aac Asn	cgc Arg	atc Ile	tac Tyr 75	cgc Arg	atc Ile	ctg Leu	gag Glu	ctg Leu 80		240
aac Asn	ggc Gly	tac Tyr	gaa Glu	cct Pro 85	gcc Ala	tac Tyr	gcc Ala	ggc Gly	tcc Ser 90	gtc Val	ttt Phe	ctc Leu	ggc Gly	tgg Trp 95	gcc Ala		288
cag Gln	aaa Lys	agg Arg	ttc Phe 100	Gly	aag Lys	cgc Arg	aac Asn	acc Thr 105	atc Ile	tgg Trp	ctg Leu	ttt Phe	ggg Gly 110	ccg Pro	gcc Ala		336
acc Thr	acg Thr	ggc Gly 115	Lys	acc Thr	aac Asn	atc Ile	gcg Ala 120	gaa Glu	gcc Ala	atc Ile	gcc Ala	cac His 125	gcc Ala	gtg Val	ccc Pro		384
ttc Phe	tac Tyr 130	Gly	tgc Cys	gtc Val	aac Asn	tgg Trp 135	acc Thr	aat Asn	gag Glu	aac Asn	ttt Phe 140		ttc Phe	aat Asn	gat Asp		432
tgc Cys 145	gtc Val		aag Lys	atg Met	gtg Val 150	atc Ile	tgg Trp	tgg Trp	gag Glu	gag Glu 155	ggc	aag Lys	atg Met	acg Thr	gcc Ala 160		480
		gtg Val	gag Glu	tcc Ser 165	Ala	aag Lys	gcc Ala	att Ile	ctc Leu 170	GTY	ggc Gly	ago Ser	aag Lys	gtg Val 175	cgc Arg		528
gtg Val	gac Asp	caa Glr	a aag Lys 180	Cys	aag Lys	tcg Ser	tcc Ser	gcc Ala 185	GIII	atc	gac Asp	ccc Pro	acc Thr 190		gtg Val		576
atc Ile	gtc Val	aco Thi	r Sei	aac Asr	acc Thr	aac Asn	atg Met 200	. Cys	gcc Ala	gtg Val	att Ile	gac Asp 205		aa Ası	agc n Ser		624
acc Thr	acc Thr	Phe	c gaç e Glu	g cad 1 His	cag s Glr	cag Glr 215	Pro	tto Lev	g caç ı Glr	g gac n Asp	220	,	g tto t Phe	aaa E Ly	a ttt s Phe		672
gaa Gl: 225	ı Leı	c äc ı Th	c cg r Ar	c cg	t cto g Lei 230	ı Gil	g cat ı His	gao S Asp	c ttt p Phe	ggc Gly 235	y шy.	g gte s Va	g aca l Thi	a aa r Ly	g cag s Gln 240		720
ga: Gl:	a gto u Val	c aa l Ly	a ga s Gl	g tt u Ph	e Phe	c cgo	c tgg g Tr	g gc	g cad a Gli 25	i vol	t ca p Hi	c gt s Va	g ac	c ga r Gl 25	g gtg u Val 5		.768
gc	g ca a Hi	t ga s Gl	g tt u Ph 26	е Ту	c gto r Vai	c aga	a aad g Ly:	g gg s Gl 26	ğ Gπ	a gc	c aa a As	c aa n Ly	a ag s Ar 27	9	c gcc o Ala	:	816
cc Pr	c ga o As	t ga p As 27	p Al	g ga a As	t aa p Ly	a ag s Se	c ga r Gl 28	u PL	c aa o Ly	g cg s Ar	g gc g Al	c tg a Cy 28		c to	a gto er Val	:	864

gcg Ala	gat Asp 290	cca Pro	tcg Ser	acg Thr	ser	gac o Asp i 295	gcg Ala	gaa Glu	gga g Gly A	gct (	ccg Pro 300	gtg Val	gac Asp	ttt Phe	: ge : A.	cc la	912
gac Asp 305		tac Tyr	caa Gln	aac Asn	aaa Lys 310	tgt Cys	tct Ser	cgt Arg	1110	gcg Ala 315	ggc Gly	atg Met	ctt Leu	caç Glr	g a n M 3	tg et 20	960
ctg Leu	ttt Phe	ccc Pro	tgc Cys	aag Lys 325	aca Thr	tgc Cys	gag Glu	aga Arg	atg Met 330	aat Asn	cag Gln	aat Asn	ttc Phe	aac Asi 33!	c a n I 5	tt le	1008
tgc Cys	ttc Phe	acg Thr	cac His 340	ggg Gly	acg Thr	aga Arg	gac Asp	tgt Cys 345	tca Ser	gag Glu	tgc Cys	ttc Phe	ccc Pro 350	gg <sub>g</sub>	у У	tg al	1056
tca Ser	gaa Glu	tct Ser 355	GIn	ccg Pro	gtc Val	gtc Val	aga Arg 360	aag Lys	agg Arg	acg Thr	tat Tyr	cgg Arg 365	aaa Lys	ct Le	c t	gt Cys	1104
gco Ala	att Ile 370	His	cat His	ctg Leu	ctg Leu	ggg Gly 375	cgg Arg	gct Ala	ccc Pro	gag Glu	att Ile 380	gct Ala	tgc Cys	tc Se	g (	gcc Ala	1152
tgo Cys 385	Asp	cto Lev	g gto ı Val	: aac . Asn	gtg Val 390	gac Asp	ctg Leu	gat Asp	gac Asp	tgt Cys 395		tct Ser	gag Glu	g ca 1 Gl	na In	taa	1200
<2: <2:	10> 11> 12> 13>	9 399 PRT AAV	-1													٠.	
<2: <2: <2: <4 Me	11> 12> 13>	399 PRT AAV		L Gl <sub>3</sub> 5	y Trp	o Leu	Val	. Asp	Arg 10	Gly	, Ile	e Thi	r Se	r G:	lu 5	Lys	
<2: <2: <2: <4 Me 1	11> 12> 13> 00> t Gl	399 PRT AAV 9 u Le	u Vai	5		o Leu o Glm			10								
<2: <2: <2: <4 Me 1 Gl	11> 12> 13> 00> t Glu	399 PRT AAV 9 u Le p Il n Se 35	u Vale Gl 20 r Ar	5 n Gli g Se:	ı Asp	o Glm	Ala Ly: 40	Ser 25 s Ala	Tyr A Ala	: Ile	e Se	r Pho p As 45	e As 30 n Al	n A a G	la ly	Ala Lys	
<2: <2: <2: <4 Me 1 G1	11> 12> 13> 00> t Glu n Tr; r As	9 u Le p II n Se 35	u Va. e Gl 20 r Ar	5 n Glu g Se: u Th	ı Asp r Gli r Ly	o Glm n Ile s Sei 55	Ala Ly: 40	Ser 25 S Ala	Tyr  A Ala	: Ile	e Se u As r Le 60	r Pho P As 45 u Va	e As 30 n Al 1 Gl	n A a G y P	la ly 'ro	Ala Lys Ala	
<2: <2: <2: <4 Me 1 G1 Se I1	11> 12> 13> 00> t Glu n Tr; r As	9 Le p Il n Se 35	u Vale Gli 20 r Ar	n Glu g Se u Th	ı Asp r Glı r Ly e Ly 70	o Glm n Ile s Sen 55	Ala Ly: 40 Ala r As	Ser 25 S Ala a Pro	Tyr Ala Ala	Lev Ty: Ty:	e Se.  As:  Le  60  r Ar	r Ph P As 45 u Va	e As 30 n Al 1 Gl	n A. a G .y P	la ly ro	Ala Lys Ala Leu 80	
<2: <2: <2: <4 Me 1 Gl Se II P: 6:	11> 12> 13> 00> t Glu n Tr r As e Me - 50 ro Pr	9 Le p Il n Se 35 t Al	u Vale Gli 20 r Ar a Le	n Glug Sei u Th p Il	r Glir Ly e Ly 70	s Sei 55 s Th:	Ala Ly: 40 Ala r As	a Ser 25 Ala Pro n Ar	Tyr Ala Ala S Asp	Let Ty: 75	r Le 60 r Ar	r Phop As 45 u Va	e As 30 n Al 1 Gl	n A. a G y P eu G	la ly ro Glu	Ala Lys Ala Leu 80 Ala	
<2: <2: <2: <4 Me 1 Gl Se II P: 6:	11> 12> 13> 00> t Glu n Tr r As e Me - 50 ro Pr	9 Le p Il n Se 35 t Al	u Vale Gli 20 r Ar a Le	n Glug Sei u Th p Il	r Glir Ly e Ly 70	o Glm n Ile s Sen 55	Ala Ly: 40 Ala r As	a Ser 25 Ala Pro n Ar	Tyr A Ala A Asp G Ile Y Se 90 r Il	Let Ty: 75	r Le 60 r Ar	r Phop As 45 u Va	e As 30 n Al 1 Gl	n A. a G y P eu G	la ly ro Glu	Ala Lys Ala Leu 80 Ala	
<2: <2: <2: <4 Me 1 G1 Se II Pr 6: A: G: G: G: G: C:	11> 12> 13> 00> t Glu n Tr r As e Me - 50 ro Pr in Glu ln Ly	9 Le p Il n Se 35 ct Al ly Ty	u Vale Gli 20 r Ar a Le a As	n Glug Seiu Th	r Glir Ly e Ly 70 o Al	s Sei 55 s Th:	Ala Ly: 40 Ala r As r Ala	a Ser 25 s Ala a Pro n Ar a Gl	Tyr Ala Ala Asp Se 90 r Il	E Ile Lev Ty: 75 r Va	r Le 60 r Ar	r Phop As 45 u Value II Le Le Le Le Le II la H:	e As 30 n Al l Gl e Le Eu Gl	n A  a G  y P  Lly T  cly I  lly I	la ly Pro Glu Pro	Ala Lys Ala Leu 80 Ala	

Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 145 150 150 155

Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 165 170 175

Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val 180 185 190

Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 195 200 205

Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe 210 220

Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln 225 230 235

Glu Val Lys Glu Phe Phe Arg Trp Ala Gln Asp His Val Thr Glu Val 245 250 255

Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Asn Lys Arg Pro Ala 260 265 270

Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val 275 280 285

Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala 290 295 300

Asp Arg Tyr Gln Asn Lys Cys Ser Arg His Ala Gly Met Leu Gln Met 305 315 320

Leu Phe Pro Cys Lys Thr Cys Glu Arg Met Asn Gln Asn Phe Asn Ile 325 330 335

Cys Phe Thr His Gly Thr Arg Asp Cys Ser Glu Cys Phe Pro Gly Val 340 345

Ser Glu Ser Gln Pro Val Val Arg Lys Arg Thr Tyr Arg Lys Leu Cys 355 360 365

Ala Ile His His Leu Leu Gly Arg Ala Pro Glu Ile Ala Cys Ser Ala 370 ° 380

Cys Asp Leu Val Asn Val Asp Leu Asp Asp Cys Val Ser Glu Gln 385 395

<210> 10

<211> 969

<212> DNA

<213> AAV-1

<220>

<221> CDS

<222> (1)..(966)

<223>

<220> misc\_feature <221> (943)..(944) <222> minor splice site <223> <400> 10 atg gag ctg gtc ggg tgg ctg gtg gac cgg ggc atc acc tcc gag aag 48 Met Glu Leu Val Gly Trp Leu Val Asp Arg Gly Ile Thr Ser Glu Lys cag tgg atc cag gag gac cag gcc tcg tac atc tcc ttc aac gcc gct 96 Gln Trp Ile Gln Glu Asp Gln Ala Ser Tyr Ile Ser Phe Asn Ala Ala tcc aac tcg cgg tcc cag atc aag gcc gct ctg gac aat gcc ggc aag 144 Ser Asn Ser Arg Ser Gln Ile Lys Ala Ala Leu Asp Asn Ala Gly Lys 40 35 atc atg gcg ctg acc aaa tcc gcg ccc gac tac ctg gta ggc ccc gct 192 Ile Met Ala Leu Thr Lys Ser Ala Pro Asp Tyr Leu Val Gly Pro Ala 50 ccg ccc gcg gac att aaa acc aac cgc atc tac cgc atc ctg gag ctg 240 Pro Pro Ala Asp Ile Lys Thr Asn Arg Ile Tyr Arg Ile Leu Glu Leu aac ggc tac gaa cet gec tac gec ggc tec gtc ttt etc ggc tgg gec 288 Asn Gly Tyr Glu Pro Ala Tyr Ala Gly Ser Val Phe Leu Gly Trp Ala 90 85 cag aaa agg ttc ggg aag cgc aac acc atc tgg ctg ttt ggg ccg gcc 336 Gln Lys Arg Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala 100 acc acg ggc aag acc aac atc gcg gaa gcc atc gcc cac gcc gtg ccc 384 Thr Thr Gly Lys Thr Asn Ile Ála Glu Ála Ile Ála His Ála Val Pro 125 120 115 ttc tac ggc tgc gtc aac tgg acc aat gag aac ttt ccc ttc aat gat 432 Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp tgc gtc gac aag atg gtg atc tgg tgg gag gag ggc aag atg acg gcc 480 Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 155 150 aag gtc gtg gag tcc gcc aag gcc att ctc ggc ggc agc aag gtg cgc 528 Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 170 165 gtg gac caa aag tgc aag tcg tcc gcc cag atc gac ccc acc ccc gtg 576 Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val atc gtc acc tcc aac acc aac atg tgc gcc gtg att gac ggg aac agc 624 Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 200 acc acc ttc gag cac cag cag ccg ttg cag gac cgg atg ttc aaa ttt 672

		210					215												
	gaa Glu 225	ctc Leu	acc Thr	cgc Arg	cgt Arg	ctg Leu 230	gag Glu	cat His	gac Asp	ttt Phe	ggc Gly 235	aag Lys	gtg Val	ac Th	a a	aag Lys	cag Gln 240		720
	gaa Glu	gtc Val	aaa Lys	gag Glu	ttc Phe 245	ttc Phe	cgc Arg	tgg Trp	gcg Ala	cag Gln 250	110 P	cac	gto Val	g ac	cc (	gag Glu 255	gtg Val		768
	gcg Ala	cat His	gag Glu	ttc Phe 260	Tyr	gtc Val	aga Arg	aag Lys	ggt Gly 265	gga Gly	gcc Ala	aac Asr	aaa Ly:	a a a s A . 2	ga rg 70	ccc Pro	gcc Ala		816
	ccc Pro	gat Asp	gac Asp 275	Ala	gat Asp	aaa Lys	agc Ser	gag Glu 280	ccc Pro	aag Lys	cgg Arg	g Ala	tg a Cy 28	с с s Р 5	cc ro	tca Ser	gtc Val		864
	gcg Ala	gat Asp 290	cca Pro		acg Thr	tca Ser	gac Asp 295	Ala	gaa Glu	gga Gly	gct Ala	2 CÇ	-	g g l A	ac .sp	ttt Phe	gcc Ala		912
	gac Asp 305	agg Arg		ggc Gly	tgc Cys	cga Arg	tgg Trp	tta Leu	tct Ser	tco Ser	aga Arq 315	<i>–</i> –	g gc u Al	t c a A	ga	gga Gly	caa Gln 320	1	960
	cct		: tga																969
* .	<21 <21 <21 <21	.1> .2>	11 322 PRT AAV-	-1															
	<22	20> 21> 22> 23>	(94	c_fea 3) or s	(944	e ) e si	te											;	
٠.	1		• •		5		p Le												
	Gl	n Tr	p Il	e Gl 20	n Gl	u As	p Gl	n Al	a Se 25	r Ty	r Il	Le S	er P	he	Asr 30	Al	a Al	a	
			35	1			n Il	40											
:	Il	e Me		a Le	eu Th	ır Ly	/s Se 55	r Al	a Pr	o As	ър Т	yr L 6	eu V O	al'	Gl	y Pi	:o A]	_a	
	Pr 65		o Al	a As	sp II	le Ly 70	ys Tł O	nr As	ıA n	g I	le T	yr A 5	rg I	le	Le	u G	Lu Le 80	eu O	
			Ly Ty	yr G	Lu P:	ro Al	la T	yr Al	La G	Ly S	er V O	al F	he l	Leu	Gl	у Т. 9	rp A	la	

- Gln Lys Arg Phe Gly Lys Arg Asn Thr Ile Trp Leu Phe Gly Pro Ala
  100 105 110
- Thr Thr Gly Lys Thr Asn Ile Ala Glu Ala Ile Ala His Ala Val Pro 115 120 125
- Phe Tyr Gly Cys Val Asn Trp Thr Asn Glu Asn Phe Pro Phe Asn Asp 130
- Cys Val Asp Lys Met Val Ile Trp Trp Glu Glu Gly Lys Met Thr Ala 145 150 150
- Lys Val Val Glu Ser Ala Lys Ala Ile Leu Gly Gly Ser Lys Val Arg 165 170 175
- Val Asp Gln Lys Cys Lys Ser Ser Ala Gln Ile Asp Pro Thr Pro Val 180 185
- Ile Val Thr Ser Asn Thr Asn Met Cys Ala Val Ile Asp Gly Asn Ser 195 200 205
- Thr Thr Phe Glu His Gln Gln Pro Leu Gln Asp Arg Met Phe Lys Phe 210 220
- Glu Leu Thr Arg Arg Leu Glu His Asp Phe Gly Lys Val Thr Lys Gln 225 235 240
- Glu Val Lys Glu Phe Phe Arg Trp Ala Gln Asp His Val Thr Glu Val 245
- Ala His Glu Phe Tyr Val Arg Lys Gly Gly Ala Asn Lys Arg Pro Ala 260 265 270
- Pro Asp Asp Ala Asp Lys Ser Glu Pro Lys Arg Ala Cys Pro Ser Val 275 280 285
- Ala Asp Pro Ser Thr Ser Asp Ala Glu Gly Ala Pro Val Asp Phe Ala 290 295 300
- Asp Arg Tyr Gly Cys Arg Trp Leu Ser Ser Arg Leu Ala Arg Gly Gln 305 310

Pro Leu

<210> 12

<211> 2211 <212> DNA

<213> AAV-1

<220>

<221> CDS

<222> (1)..(2208)

<223>

<400> 12 atg gct gcc gat ggt tat ctt cca gat tgg ctc gag gac aac ctc tct Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser

																_	816
	Lys	Gln	11e 260	Ser	Ser.	Ala	ser	265	OLY				270				010
tac Tyr	ttc Phe	ggc Gly 275	tac Tyr	agc Ser	acc Thr	ccc Pro	tgg Trp 280	Gly ggg	tat Tyr	ttt Phe	gat Asp	ttc Phe 285	aac Asn	aga Arg	tt. Ph	c e	864
cac His	tgc Cys 290	cac His	ttt Phe	tca Ser	cca Pro	cgt Arg 295	gac Asp	tgg Trp	cag Gln	cga Arg	ctc Leu 300		aac Asn	aac Asn	aa As	t n	912
tgg Trp 305		ttc Phe	cgg Arg	ccc Pro	aag Lys 310	aga Arg	ctc Leu	aac Asn	ttc Phe	aaa Lys 315		ttc Phe	aac Asn	atc Ile	ca Gl 32	a .n !0	960
	aag Lys	gag Glu	gtc Val	acg Thr 325	acg Thr	aat Asn	gat Asp	ggc Gly	gtc Val 330	aca Thr	acc Thr	ato Ile	gct Ala	aat Asn 335	aa As	ic sn	1008
ctt Leu	acc Thr	agc Ser	acg Thr	Val	caa Gln	gtc Val	ttc Phe	tcg Ser 345	TOP	tcg Ser	gag Glu	g tad 1 Tyl	c cag c Gln 350	r ctt Leu	co 1 Pi	eg co	1056
tac Tyr	gtc Val	ctc Leu 355	ggc Gly		gcg Ala	cac His	cag Gln 360	, Gry	tgc Cys	cto Lev	cot Pro	5 CC 5 Pro 36	g tto Phe	cco Pro	g g( 5 A.	cg la	1104
gac Asp	gtg Val	tto Phe		att : Ile	ccg Pro	caa Gln 375	TÀT	: Gly	tac Tyr	cto	g aco 1 Th 38		c aad u Asi	aat n Asi	c gr n G	gc ly	1152
ago Ser 385	caa Glr		c gto a Val	r Gl <sup>7</sup> d dds	a cgt 7 Arg 390	, ser	tco Sei	ttt Phe	tac Tyı	tg Cy: 39		g ga u Gl	a ta u Ty	t tt r Ph	с с е Р 4	ct ro 00	1200
		g ato n Me	g cto	g aga u Arg	a aco		aa Ası	c aaq n Ası	c tti n Phe		c tt r Ph	c ag e Se	c ta r Ty	c ac r Th 41	c t r P 5	tt he	1248
ga Gl	g ga u Gl	a gt u Va	g cc 1 Pr 42	t tt		c ago	c ag r Se	c ta r Ty 42	T 1	g ca a Hi	c aç s Se	ıc ca er Gl	ig ag n Se 43	c ct r Le 0	g g u A	lac Asp	1296
cg Ar	g ct g Le		g aa t As		t ct o Le	c at u Il	c ga e As 44	P OT	a ta n Ty	c ct r Le	g ta u Ty		ac ct yr Le 45	g aa u As	ic a	aga Arg	1344
ac Th	t ca r Gl 45	a aa n As		g to n Se	c gg r Gl	a ag y Se 45	T VT	c ca .a Gl	a aa n As	c aa n Ly	ag ga ys A: 4	ac t' sp L	tg ct eu Le	ig tt eu Pl	it a	agc Ser	1392
cg Ar 46	rt gg g Gl		et co er Pr	a go co Al	t gg a Gl 47	.у ме	g to	et gt er Va	t ca al Gl		cc a ro L 75	aa a ys A	ac to sn T	gg c	ta eu	cct Pro 480	1440
		cc to	gt ta ys Ty	at co yr Ai	gg ca	ag ca In Gl	ag co in A:	gc gt rg Va	tt to al Se	ct a er L	aa a ys T	ca a hr L	aa a ys T	ca g hr A	ac sp	aac. Asn	1488

403	aat 1536
aac aac agc aat ttt acc tgg act ggt gct tca aaa tat aac ctc a Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu A 500 505	
ggg cgt gaa tcc atc atc aac cct ggc act gct atg gcc tca cac ac Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His I 515	aaa 1584 Lys
gac gac gaa gac aag ttc ttt ccc atg agc ggt gtc atg att ttt case atg agc ggt gtc atg att ttt gas Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe 530 535	gga 1632 Gly
aaa gag agc gcc gga gct tca aac act gca ttg gac aat gtc atg Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met 545 550 555	att 1680 Ile 560
aca gac gaa gag gaa att aaa gcc act aac cct gtg gcc acc gaa Thr Asp Glu Glu Glu Lys Ala Thr Asn Pro Val Ala Thr Glu 575	aga 1728 Arg
ttt ggg acc gtg gca gtc aat ttc cag agc agc agc aca gac cct Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro 580 585	gcg 1776 Ala
acc gga gat gtg cat gct atg gga gca tta cct ggc atg gtg tgg Thr Gly Asp Val His Ala Met Gly Ala Leu Pro Gly Met Val Trp 605	caa 1824 Gln
gat aga gac gtg tac ctg cag ggt ccc att tgg gcc aaa att cct Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro 615	cac 1872 His
aca gat gga cac ttt cac ccg tct cct ctt atg ggc ggc ttt gga Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly 625 630 635	ctc 1920 Leu 640
aag aac ccg cct cct cag atc ctc atc aaa aac acg cct gtt cct Lys Asn Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro 645 650	gcg 1968 Ala 6
aat cct ccg gcg gag ttt tca gct aca aag ttt gct tca ttc atc Asn Pro Pro Ala Glu Phe Ser Ala Thr Lys Phe Ala Ser Phe Ile 660 665	c acc 2016 e Thr
caa tac tcc aca gga caa gtg agt gtg gaa att gaa tgg gag ctg Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Lev 675 680	g cag 2064 u Gln
aaa gaa aac agc aag cgc tgg aat ccc gaa gtg cag tac aca tc Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Val Gln Tyr Thr Se 690 695	c aat 2112 r Asn
tat gca aaa tot gcc aac gtt gat ttt act gtg gac aac aat gg Tyr Ala Lys Ser Ala Asn Val Asp Phe Thr Val Asp Asn Asn Gl 705	a ctt 2160 y Leu 720
tat act gag cct cgc ccc att ggc acc cgt tac ctt acc cgt cc Tyr Thr Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Pr	cc ctg 2208 co Leu

725

730

<210>	13
<211>	736
<212>	PRT
/213>	AAV-1

<400> 13
Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro 50 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro 115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg 130 135

Pro Val Glu Gln Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly 145 150 150

Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln Thr 165 170 175

Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro Pro 180 185 190

Ala Thr Pro Ala Ala Val Gly Pro Thr Thr Met Ala Ser Gly Gly Gly 195 200 205

Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn Ala 210 215 220

Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val Ile 225 230 235 240

Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His Leu 245 250 250

Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly Ala Ser Asn Asp Asn His

 Tyr
 Phe
 Gly 275
 Tyr
 Pro 280
 Tyr
 Gly Tyr
 Phe
 Asn Arg
 Phe
 Asn Asn Asn Asn Asn Asn 296
 Phe 290
 Arg 295
 Asp Trp Gln
 Arg Leu 11e Asn Asn Asn Asn 300
 Asn Asn Asn Asn 300
 Asn Asn Asn Asn 320
 Asn 310
 Arg Sp Trp 310

450 455 460

Arg Gly Ser Pro Ala Gly Met Ser Val Gln Pro Lys Asn Trp Leu Pro

Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Lys Thr Asp Asn 490 495

Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala Ser Lys Tyr Asn Leu Asn 500 505

Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr Ala Met Ala Ser His Lys 515 520 525

Asp Asp Glu Asp Lys Phe Phe Pro Met Ser Gly Val Met Ile Phe Gly 530

Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala Leu Asp Asn Val Met Ile 545 550 555 560

Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn Pro Val Ala Thr Glu Arg 565 570 575

Phe Gly Thr Val Ala Val Asn Phe Gln Ser Ser Ser Thr Asp Pro Ala

		595			Ala		000													
	610				Leu	613														
625					His 630					-	_									
Lys	Asn	Pro	Pro	Pro 645	Gln	Ile	Leu	Ile	Lys 650	s As )	sn	Thr	Pr	o V	al	Pr 65	o <i>P</i> 5	Ala		
Asn	Pro	Pro	Ala 660	Glu	Phe	Ser	Ala	Thr 665	Lys	s Pl	he	Ala	Se	r E	he 570	Il	e T	rhr		
Gln	Tyr	Ser 675	Thr	Gly	Gln	Val	Se1	: Val	_ Gl	u I	le	Glu	Tr 68	p (	3lu	Le	u (	Gln		
Lys	Glu 690	Asn	ser	Lys	Arg	Trp 695	Ası	n Pro	o Gl	u V	al	Gĺn 700	Τ	r'	Гhr	Se	er 2	Asn		
Туг 705		Lys	s Ser	Ala	Asn 710	Val	. Ası	Phe	e Th	r V 7	al 15	Asp	As	sn I	Asn	G]	-У	Leu 720		
Туг	Thr	: Glu	ı Pro	729	g Pro	Ile	e Gl	y Th	r Ar 73	g T	'yr	Leu	ı Tl	nr.	Arg	73	5 35	Leu		
<21 <21 <22 <22 <22	11> 12> 13> 20> 21> 22>	14 1800 DNA AAV: CDS (1)		797)																
	23> 400>	14													<b>6</b> 2.		. a. a	cca		48
ac Th 1	g gc r Al	t cc a Pr	t gg	a aa y Ly 5	g aa s Ly	a cg s Ar	t co	g gt o Va	a g al G	_ ~	caç Glr	g to 1 Se	g c r E	ro	Gl	n G	1u .5	Pro		
	c to p Se	c to r Se	c to er Se 20	r Gi	ıc at .y Il	c gg e Gl	y L	ag ac ys Tl 25		gc ly	caç Gl:	g ca n Gl	g c n I	ccc Pro	gc Al 30	t a a I	aa Lys	aaç Lys	J 5	96
aç Ar	ja ct :g Le	c aa eu As	sn Pr	t gg ne Gl	gt ca Ly Gl	g ac n Th	et go ar G	TA V	ac t sp S	ca	ga Gl	g to u S∈	er V	gtc Val 45	cc Pr	c (	gat Asp	CCa Pro	a o	144
ca G1	aa co Ln Pi	co L	tc go eu Gi	ga ga Ly Gi	aa co lu Pi	t co to P:	EO A	ca a la T	cc c hr E	cc Pro	gc Al	t go a Al 60	ct La O	gtg Val	G1	ıa Y	cct Pro	ac Th	t r	192
ae Ti 6:	hr M	tg g et A	ct to la S	ca g er G	gc gg ly G	ry G	gc g ly A	ca c la P	ca a	atg Met	gc Al 75		ac sp	aat Asn	: aa n As	ac sn	gaa Glu	a gg ı Gl 80	С. У	240

gcc (	Asp	Gly	Val	61y 85	Asn A	чта	Ser	GLY	90			•	_	95				288
tgg (	Leu	Gly	Asp 100	Arg	Val	TIE	IIII	105	501		9		110					336
ccc Pro	acc Thr	tac Tyr 115	aat Asn	aac Asn	cac His	ctc Leu	tac Tyr 120	aag Lys	caa Gln	atc Ile	tcc Ser	agt Ser 125	gct Ala	tca Se:	aa r T	cg hr		384
gly ggg	gcc Ala 130	agc Ser	aac Asn	gac Asp	aac Asn	cac His 135	tac Tyr	ttc Phe	ggc Gly	tac Tyr	agc Ser 140		e ccc	tg Tr	p C	ej A 188		432
tat Tyr 145	ttt Phe	gat Asp	ttc Phe	aac Asn	aga Arg 150	ttc Phe	cac His	tgc Cys	cac His	ttt Phe 155		Pro	a cgt o Arg	ga JAs	c t p :	rgg Frp 160		480
cag Gln	cga Arg	ctc Leu	ato	aac Asn 165	aac Asn	aat Asn	tgg Trp	gga Gly	ttc Phe 170	11. 9	ccc	aaq Ly:	g aga s Arg	a ct g Le 17	c u 2	aac Asn		528
ttc Phe	aaa Lys	cto	tto Phe	. Asn	atc Ile	caa Gln	gtc Val	aag Lys 185	Gra	gtc Val	aco Thi	g ac	g aa r Asi 19	t ga n As O	t p	ggc Gly		576
gtc Val	aca Thr	acc Thr 195	: Ile	gct Ala	aat Asn	aac Asn	ctt Leu 200	rint	agc Ser	acg Thr	ggtt Val	ca L Gl 20	a gt n Va 5	c tt l Ph	ic ne	tcg Ser		624
gac Asp	tcg Ser 210	Glu	g tac ı Ty	cag Glr	g ctt Leu	ccg Pro 215	TAT	gtc Val	cto Lev	. ggc	z tci 7 Se: 22		g ca a Hi	c ca s G	ag Ln	ggc Gly		672
tgc Cys 225	Leu	cci	t cc	g tto o Phe	c ccg e Pro 230	Ala	gad Asp	gto Val	g tto L Phe	ato Met 23		t co e Pr	g ca o Gl	a tan	ac yr	ggc Gly 240		720
		g ac	g ct r Le	c aac u Ası 24	c aat n Asn 5	ggo Gly	c ago y Se:	c caa r Gli	a gco n Ala 25		g gg 1 Gl	a co y Ai	gt to :g Se		cc er 55	ttt		768
tac Tyr	tg:	 c ct s Le	g ga u Gl 26	u Ty	t tto r Phe	e Pr	t tc o Se	t ca r Gl: 26	11 110	g ct t Le	g ag u Ar	a ad g Tl	eg gg hr Gl 2	gc a Ly A 70	ac	aac Asr	: 1	816
ttt Phe	ac Th	c tt r Ph 27	e Se	c ta r Ty	c acc	c tt r Ph	t ga e Gl 28	u Gi	a gt u Va	g cc l Pr	t tt o Ph		ac a is S 85	gc a er S	igc Ser	tac Ty	= r	864
gc: Ala	g ca a Hi 29	s Se	ıc ca er Gl	ig ag .n Se	c cte	g ga u As 29	p Ar	ıg ct :g Le	g at u Me	g aa t As		ct c ro L	tc a eu I	tc q le <i>P</i>	gac Asp	Gl	a n	912
ta Ty 30	r Le	g ta u Ty	at ta yr Ty	ac ct yr Le	g aa eu As 31	n Ar	ga ac :g Th	et ca nr Gl	aa aa Ln As	,11 02	ag te ln S	cc g er G	ga a Sly S	gt o er i	gcc Ala	: ca Gl 32	a n 0	960

	Asn	Lys	Asp		Leu 1 325	rne .	ser /	ALG	GLY .	330			-		335			1008
	cag Gln	ccc Pro	aaa Lys	aac Asn 340	tgg Trp	cta Leu	cct Pro	gga Gly	ccc Pro 345	tgt Cys	tat Tyr	cgg Arg	cag Gln	cag Gln 350	cgc Arg	gtt Val		1056
	tct Ser	aaa Lys	aca Thr 355	aaa Lys	aca Thr	gac Asp	aac Asn	aac Asn 360	aac Asn	agc Ser	aat Asn	ttt Phe	acc Thr 365	tgg Trp	act Thr	ggt Gly		1104
٠.	gct Ala	tca Ser 370	Lys	tat Tyr	aac Asn	ctc Leu	aat Asn 375	Gly ggg	cgt Arg	gaa Glu	tcc Ser	atc Ile 380	atc Ile	aac Asn	cct Pro	Gl?	; ,	1152
	act Thr 385	Ala	atg Met	gcc Ala	tca Ser	cac His 390	aaa Lys	gac Asp	gac Asp	gaa Glu	gac Asp 395	aag Lys	ttc Phe	ttt Phe	ccc Pro	ato Met 400	) - )	1200
			gtc Val	atg Met	att Ile 405	ttt Phe	gga Gly	aaa Lys	gag Glu	agc Ser 410	gcc Ala	gga Gly	gct Ala	tca Ser	aac : Asr 415	act Th:	ī.	1248
	gca Ala	ttg Leu	g gac . Asp	aat Asn 420	Val	atg Met	att Ile	aca Thr	gac Asp 425	gaa Glu	gag Glu	gaa Glu	att	aaa Lys 430	a gco s Ala	a Th	t r	1296
	aac Asn	cct Pro	gto Val	gcc Ala	acc Thr	gaa Glu	aga Arg	ttt Phe 440	Gry	acc Thr	gtg Val	gca Ala	gto Val 445	aat Ası	tte n Phe	c ca e Gl	g n	1344
	ago Ser	ago Sei 450	r Sei	aca Thr	gac Asp	cct Pro	gcg Ala 455	IIIT	gga Gly	gat Asp	gtg Val	cat His 460		t ato	g gg	a gc y Al	a a	1392
	tta Lev 465	ı Pr	t ggo o Gl	c atg y Met	gtg Val	tgg Trp 470	GIII	gat Asp	aga Arg	gac Asp	gto Val	1	c ct	g ca u Gl	g gg n Gl	t cc y Pr 48	c 0 10	1440
			g gc p Al	c aaa a Lys	a att s Ile 485	Pro	cac His	aca Thi	a gat r Asp	gga Gly 490		tt s Ph	t ca e Hi	c cc s Pr	g to o Se 49	t co r Pr	t o	1488
	ct Le	t at u Me	g gg t Gl	c gg y Gl	y Phe	gga Gly	tet Let	aaq Ly:	g aad s Asr 505		g cc o Pr	t cc o Pr	t ca o Gl	g at n Il 51	c ct e Le .0	c at u II	c Le	1536
	aa Ly	a aa s As	c ac n Th	g cc r Pr	t gti o Val	cci L Pro	z gcg o Ala	g aa a As: 52	II ET	t cc D Pr	g gc o Al	g ga a Gl	g tt u Ph 52		er Al	ct a La T	ca hr	1584
	aa Ly	s Ph	t go ne Al	t tc a Se	a tte r Ph	c at	c ac e Th 53	E GT	a ta n Ty	c tc r Se	c ac r Th	a gg ir Gl 54	- 1	aa gi Ln Va	tg ad al S	gt g er V	tg al	1632
	G]			aa tg Lu Tr	g ga p Gl	g ct u Le 55	u Gi	g aa n Ly	a ga 's Gl	a aa u As		jc aa er Ly 55	ag co ys A	gc t rg T	gg a rp A	at c sn P 5	ro 60	1680

· .	gaa gi Glu Va	tg ca	ag ta ln T	yr T	ca t hr S 65	cc a er A	at t sn T	at g yr A		aa t ys S 70	ct g Ser A	cc a la A	ac g sn V	tt g al A	at t sp P	tt he	1728
	act g Thr V	tg g al A	sp A	ac a sn A 80	at g sn G	ga c ly L	tt t eu T	Y	ct g hr 6	ag o	ect o Pro P	gc c rg P	cc a ro I	tt g le 0	ggc a Sly T	icc hr	1776
	cgt t Arg T	yr L	tt a eu T 95	cc c	gt c Arg P	cc c	tg t eu	aa									1800
	<210><211><211><212><213>	> 59	9														
÷ .	<400> Thr F	> 15 Ala B	o Pro (	Gly :	Lys : 5	Lys .	Arg	Pro	Val	Glu 10	Gln	Ser	Pro	Gln	Glu 15	Pro	
	Asp S	Ser S	Ser :	Ser 20	Gly	Ile	Gly	Lys	Thr 25	Gly	Gln	Gln	Pro	Ala 30	Lys	Lys ,	
	Arg 1		Asn 35	Phe	Gly	Gln	Thr	Gly 40	Asp	Ser	Glu	Ser	Val 45	Pro	Asp	Pro	
	Gln	Pro :	Leu	Gly	Glu	Pro	Pro 55	Ala	Thr	Pro	Ala	Ala 60	Val	Gly	Pro	Thr	•
٠.	Thr 1		Ala	Ser	Gly	Gly 70	Gly	Ala	Pro	Met	Ala 75	Asp	Asn	Asn	Glu	Gly 80	
	Ala	Asp	Gly	Val	Gly 85	Asn	Ala	Ser	Gly	Asn 90	Trp	His	Cys	Asp	Ser 95	Thr	
	Trp	Leu	Gly	Asp 100	Arg	Val	Ile	Thr	Thr 105	Ser	Thr	Arg	Thr	Trp 110	Ala	Leu	
	Pro	Thr	Tyr 115		Asn	His	Leu	Туг 120	Lys	Gln	Ile	Ser	Ser 125	Ala	Ser	Thr	
	Gly	Ala 130		Asn	Asp	Asn	His 135	Tyr	Phe	Gly	Tyr	Ser 140	Thr	Pro	Trp	Gly	
	Tyr 145		Asp	Phe	Asn	Arg 150	Phe	His	Cys	His	Phe 155	ser	Pro	Arg	J Asp	160	
		Arg	Leu	Ile	Asn 165	Asn	Asn	Trp	Gly	Phe 170	e Arg	, Pro	Lys	. Ar	J Let 175	Asn	
	Phe	Lys	Leu	Phe	Asn		Glr	ı Val	L Lys 185	s Glu	u Val	LThr	Thi	190	n Asp 0	o Gly	·
	Val	Thr	Thr 195	: Ile		. Asr	n Ası	n Let 20	ı Thi	r Se	r Th	r Val	Gl: 20	n Val	l Ph	e Ser	:
	Asp	ser			r Glr	ı Let	ı Pr	о Ту	r Va	l Le	u Gl	y Se	r Al	a Hi	s Gl	n Gly	7

Cys Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly 225 235 240

Tyr Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe 245 250 250

Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn 260 265

Phe Thr Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr 280 285

Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln 290 295 300

Tyr Leu Tyr Tyr Leu Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln 305 310 315

Asn Lys Asp Leu Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val 325 330 335

Gln Pro Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val 340 345

Ser Lys Thr Lys Thr Asp Asn Asn Ser Asn Phe Thr Trp Thr Gly 355

Ala Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly 370 380

Thr Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro Met 385 390 395

Ser Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser Asn Thr 405 410 415

Ala Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile Lys Ala Thr 420 425 430

Asn Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln 435 440 445

Ser Ser Ser Thr Asp Pro Ala Thr Gly Asp Val His Ala Met Gly Ala 450 455 460

Leu Pro Gly Met Val Trp Gln Asp Arg Asp Val Tyr Leu Gln Gly Pro 465 470 480

Ile Trp Ala Lys Ile Pro His Thr Asp Gly His Phe His Pro Ser Pro 485 490 495

Leu Met Gly Gly Phe Gly Leu Lys Asn Pro Pro Pro Gln Ile Leu Ile 500 505

Lys Asn Thr Pro Val Pro Ala Asn Pro Pro Ala Glu Phe Ser Ala Thr 515

Lys Phe Ala Ser Phe Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val

Glu Ile	Glu	Trp G	lu Le	eu G1	n Ly	ys G	lu A	sn S	Ser 555	Lys	Arg	Trp	As	sn P	ro 60	
545 Glu Val	Gln	Tyr T			n T	yr A	la L	ys :	Ser	Ala	Asn	Va:	L A:	sp E 75	?he	
		5	65				Ŭ	, , ,								
Thr Val	Asp	Asn A 580	sn G	ly Le	eu T	yr T 5	nr 6 85	žĮu.	Pro	ALG	rro	59	0	-1		
Arg Tyr	Leu 595	Thr A	Arg P	ro L	eu											
<210> <211> <212> <213>	16 1605 DNA AAV-	1													4	
<220> <221> <222> <223>	CDS	.(160	2)													
<400> atg gc Met Al 1		ggc Gly	ggt ( Gly (	ggc g Gly <i>F</i>	gca ( Ala 1	cca Pro l	atg Met	gca Ala 10	gac Asp	aat Asr	aa n As	c ga n G	aa ( lu (	ggc Gly 15	gcc Ala	48
gac gg Asp Gl	a gtg y Val	ggt Gly 20	aat ( Asn )	gcc t Ala S	ca Ser	GTA.	aat Asn 25	tgg Trp	cat	tgo Cy:	c ga s As	t top So	cc er 0	aca Thr	tgg Trp	96
ctg gg Leu Gl	c gac y Asp 35	aga Arg	gtc Val	atc : Ile '	acc Thr	acc Thr 40	agc Ser	acc Thr	cgo	ac Th	c tg r Tr 45	g g	cc la	ttg Leu	Pro	
acc ta Thr Ty 50	ır Ası	aac Asn	cac His	ьeu	tac Tyr 55	aag Lys	caa Gln	atc Ile	tco Se:	c ag r Se 60	t go r Al	t t .a S	ca er	acg Thr	ggg	192
gcc ag Ala Se 65	gc aad er Asi	c gac n Asp	aac Asn	cac His 70	tac Tyr	ttc Phe	ggc Gly	tac Tyr	ag Se 75		c co	co T	gg 'rp	G] À	tat Tyr 80	
ttt ga Phe A	at tt sp Ph	c aac e Asn	aga Arg 85	ttc Phe	cac His	tgc Cys	cac His	ttt Phe 90	t tc e Se	a co r Pı	a co co A	gt q rg A	yac Asp	tgg Trp 95	g cag Glr	g 288 n
cga c Arg L	tc at eu Il	c aac e Asn 100	Asn	aat Asn	tgg Trp	gga Gly	ttc Phe 105		g cc g Pr	c aa o Ly	aga ys A	ga rg	ctc Leu 110	aac Asi	tte n Phe	g 336 e
aaa c Lys L	eu Ph	e Asr .5	, ite	GIII	vaı	120	010				1	25				
aca a Thr T	icc at	cc gct Le Ala	t aat a Asn	aac Asn	ctt Leu	acc Thr	ago Se:	c ac r Th	g gt	ct c al G	aa ç ln V	tc al	ttc Ph∈	: tc : Se	g ga r As	c 432 p

		130					133												
٠	Ser 145	Glu	Tyr	Gln	Leu	150	TYL	gtc Val	ьеч	GLY	155	• • -				_	160		480
	ctc Leu	cct Pro	ccg	ttc Phe	ccg Pro 165	ALA	gac <b>As</b> p	gtg Val	ttc Phe	atg Met 170	att Ile	ccg Pro	caa Glr	a ta n Ty	ac yr	ggc Gly 175	tac Tyr	•	528
	ctg Leu	acg Thr	cto Lev	aac Asn 180	Asn	ggc	agc Ser	caa Gln	gcc Ala 185	gtg Val	gga Gly	cgt Arg	tc. Se	a to r So 1	cc er 90	ttt Phe	tac Tyr		576
	tgc Cys	cto	g gaa i Gli 195	і Туі	tto Phe	cct Pro	tct Ser	cag Gln 200	Mec	ctg Leu	aga Arg	acç Thi	g gg Gl 20	са у А 5	ac .sn	aac Asn	ttt Phe		624
	acc Thr	tto Pho 21	e Se	z tac r Ty:	c acc	tti Phe	gaç Glu 219	g gaa 1 Glu 5	gtg Val	cct	tto Phe	c cad His		c a r S	gc	tac Tyr	gcg Ala		672
·	cac His	Se.	c ca r Gl:	g age n Se	c cto r Le	g ga ı Ası 23	5 AT	g ctg g Lev	atg Met	aat Asr	cct Pro 235		c at u Il	.c g .e A	jac Asp	caa Gln	tac Tyr 240		720
	cto Lev	ta 1 Ty	t ta r Ty	c ct r Le	g aa u As: 24	a Ar	a act	t caa r Glr	aat Asn	cac Glr 250		c gg r Gl	a aç y Se	jt g er A	gcc Ala	caa Gln 255	aac Asn		768
	aaç Lys	g ga s As	c tt p Le	g ct u Le 26	u Ph	t ag e Se	c cg r Ar	t ggg g Gly	tct Ser 265		a gc	t gg a Gl	c at y Me	g tet S	ser 270	gtt Val	cag Gln		816
	cco Pro	c aa o Ly	a aa s As 27	n Tr	g ct p Le	a cc u Pr	t gg o Gl	a cco y Pro 28	J Cy.	ta Ty	t cg r Ar	g ca g Gl	g ca n G: 21	ag o ln <i>l</i> 85	cgc Arg	gtt Val	tct Ser		864
	aa. Ly	a ac s Th 29	r Ly	a ac	a ga ir As	c aa p As	c aa n As 29	n As	c ago	c aa r As	t tt n Ph	t ac e Th 30	-	gg (	act Thr	ggt Gly	gct Ala	·	912
	Se	r Ly	s Ty	at aa yr As	sn Le	u As	it gg sn Gl	ıg cg .y Ar	t ga g Gl	a to u Se	c at r Il 31		c a Le A	ac .sn	cct Pro	gg Gl	c act y Thr 320	· · · · · · · · · · · · · · · · · · ·	960
	gc Al	t at .a Me	tg go	cc to la Se	er Hi	ac aa Ls Ly 25	aa ga ys As	ac ga sp As	c ga p Gl	a ga u As 33	, p	ag ti /s Pl	tc t ne P	tt	Pro	at Me 33	g ago t Sei 5	: :	1008
	G]	y V	tc a al M	et I	tt t le P 40	t g ne G	ga aa ly L	aa ga ys Gl	ig ag .u Se 34		cc gg La Gi	ga g ly A	ct t la S	ca Ser	aad Asi 350	c ac n Th O	t gca r Ala	a a	1056
	tt Le	eu A	sp A	at g sn V 55	tc a al M	tg a et I	tt a le T	ca ga hr As 30	P G	ia ga .u G.	ag ga Lu G	aa a lu I		aaa Lys 365	gc Al	c ac a Th	t aa ir As	c n	1104
	C P	ct g ro V	tg g al A	cc a la T	cc g hr G	aa a lu A	ga t rg P	tt go he G	gg ac Ly Tì	cc g nr V	tg g al A	ca g la V	tc a	aat Asn	tt Ph	c ca e Gl	ag ag In Se	c r	1152

	310															
agc Ser 385	agc Ser	aca Thr	gac Asp	cct Pro	gcg Ala 390	acc Thr	gga Gly	gat Asp	gtg Val	cat His 395	gct Ala	atg Met	gga Gly	gca Ala	tta Leu 400	1200
cct Pro	ggc Gly	atg Met	gtg Val	tgg Trp 405	caa Gln	gat Asp	aga Arg	gac Asp	gtg Val 410	tac Tyr	ctg Leu	cag Gln	ggt Gly	ccc Pro 415	att Ile	1248
tgg Trp	gcc Ala	aaa Lys	att Ile 420	cct Pro	cac His	aca Thr	gat Asp	gga Gly 425	cac His	ttt Phe	cac His	ccg Pro	tct Ser 430	cct Pro	ctt Leu	1296
atg Met	ggc Gly	ggc Gly 435	Phe	gga Gly	ctc Leu	aag Lys	aac Asn 440	ccg Pro	cct Pro	cct Pro	cag Gln	atc Ile 445	ctc Leu	atc Ile	aaa Lys	1344
aac Asn	acg Thr 450	cct Pro		cct Pro	gcg Ala	aat Asn 455	cct Pro	ccg Pro	gcg Ala	gag Glu	ttt Phe 460		gct Ala	aca Thr	aag Lys	1392
ttt Phe 465	gct Ala		ttc Phe	ato Ile	acc Thr	GIU	tac	tcc Ser	aca Thr	gga Gly 475		gtg Val	agt Ser	gtg Val	gaa Glu 480	1440
		tgg Trp	g gag o Glu	cto Leu 485	1 GIR	aaa Lys	gaa Glu	aac Asn	ago Ser 490	1-	g ego Arg	tgg Trp	aat Asr	2 CCC 1 Pro 495	gaa Glu	1488
gtg Val	cag Gln	tac Ty	aca Thi	: Sei	aat Asr	tat Tyr	gca Ala	a aaa a Lys 505	, 501	gco Ala	aac Asi	gtt Val	gat Asp 510	ttt Phe	act Thr	1536
gto Val	g gad L Asp	aac Ası 51	n Ası	gga Gl	a ctt y Lei	tat ı Tyı	act Th: 52	r GT	g cct	cgo Arg	g Pro	c atto Ile 52!		c aco y Th	c cgt r Arg	1584
tac Ty:	c cti r Lei 530	ı Th	c cg <sup>†</sup> r Ar	t cc g Pr	c cto o Le	g taa	a									 1605

<210> 17 <211> 534 <212> PRT <213> AAV-1

<400> 17
Met Ala Ser Gly Gly Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala
10 15

Asp Gly Val Gly Asn Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp 20 25 30

Leu Gly Asp Arg Val Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro 35 40 45

Thr Tyr Asn Asn His Leu Tyr Lys Gln Ile Ser Ser Ala Ser Thr Gly 50 60

- Ala Ser Asn Asp Asn His Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr 65 70 75
- Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln 85 90 95
- Arg Leu Ile Asn Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe 100 105 110
- Lys Leu Phe Asn Ile Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val
- Thr Thr Ile Ala Asn Asn Leu Thr Ser Thr Val Gln Val Phe Ser Asp 130 135
- Ser Glu Tyr Gln Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys 145 150 150
- Leu Pro Pro Phe Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr 165
- Leu Thr Leu Asn Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr 180 185 190
- Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe 195 200 205
- Thr Phe Ser Tyr Thr Phe Glu Glu Val Pro Phe His Ser Ser Tyr Ala 210 215 220
- His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr 225 230 235
- Leu Tyr Tyr Leu Asn Arg Thr Gln Asn Gln Ser Gly Ser Ala Gln Asn 255
- Lys Asp Leu Leu Phe Ser Arg Gly Ser Pro Ala Gly Met Ser Val Gln 260 265
- Pro Lys Asn Trp Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser 275 280 285
- Lys Thr Lys Thr Asp Asn Asn Asn Ser Asn Phe Thr Trp Thr Gly Ala 290 . 295
- Ser Lys Tyr Asn Leu Asn Gly Arg Glu Ser Ile Ile Asn Pro Gly Thr 305 310 315
- Ala Met Ala Ser His Lys Asp Asp Glu Asp Lys Phe Phe Pro Met Ser 325
- Gly Val Met Ile Phe Gly Lys Glu Ser Ala Gly Ala Ser Asn Thr Ala 340
- Leu Asp Asn Val Met Ile Thr Asp Glu Glu Glu Ile Lys Ala Thr Asn 365
- Pro Val Ala Thr Glu Arg Phe Gly Thr Val Ala Val Asn Phe Gln Ser 370 380

Ser 385	Ser	Thr	Asp	Pro	Ala 390	Thr	Gly	Asp	Val	His 395	Ala	Met	Gly	Ala	Leu 400	*		
Pro	Gly	Met	Val	Trp 405	Gln	Asp	Arg	Asp	Val 410	Tyr	Leu	Gln	Gly	Pro 415	Ile			
			Ile 420					120								,		
Met	Gly	Gly 435	Phe	Gly	Leu	Lys	Asn 440	Pro	Pro	Pro	Gln	Ile 445	Leu	Ile	Lys			
Asn	Thr 450	Pro	Val	Pro	Ala	Asn 455	Pro	Pro	Ala	Glu	Phe 460	Ser	Ala	Thr	Lys			
Phe		Ser	Phe	lle	Thr 470	Gln	Tyr	Ser	Thr	Gly 475	Gln	val	. Ser	Va1	Glu 480			
Ile	: Glu	ı Trp	o Glu	Leu 485	Gln	Lys	Glu	ı Asr	Ser 490	Lys	Arc	g Trp	Asn	495	Glu			
Val	Glr	тул	Th:	Ser	Asr	туг	: Ala	Lys 505	s Sei	: Ala	a Ası	n Val	510	Phe	€ Thr			
Va.	L As	o Ası 51	n Asr 5	n Gly	, Lev	туі	520	r Glu	ı Pro	Arg	g Pro	52	e Gly 5	Th:	r Arg	•		
Ту	r Le		r Ar	g Pro	) Le	1												
<2 <2	10> 11> 12> 13>	18 468 DNA AAV														\$ 		
<4 tt	aacc	18 acto	: cct	.ctct	gcg	cgct	cgct	.cg c	tcac	tgag	g co	gggc	:gacc	aaa	ggtcg	lcc	6	0
CO	acac	ccac	ı qct	ttgc	ccg	ggcg	gcct	ca g	tgag	cgag	c ga	agcgo	gcag	aga	ıgggaç	jtg	12	.0
~		rt cca	t.ca	ctac	ıaaa	ttcc	tgga	agg g	gtgg	gagto	g to	gacgt	gaat	tac	gtcat	ag	18	0
g c	.++ > /	,,,,,,,	a ata	cctat	att	agag	gtca	acg t	gagt	gttt	t go	cgaca	atttt	gc	gacac	cat	24	10
99	, <del>.</del>	999K	a ta	ratait	tta	agc	ccqa	gtg a	agcad	gcag	gg g	tctc	catt	t tg	aagcg	gga	30	00
g۱	ggt	cacy		990a.	-455	ata	caa	aat 1	ttta	cgaga	at t	gtga	ttaa	g gt	cccca	gcg	30	60
g	gttt.	gaac	g cg	state	accc	aac	attt	cta :	acaq	cttt	gt g	aact	gggt	g gc	cgaga	agg	4:	20
a	cctt	gacg	g gc	2222	3000	+c+	gaga	taa	atct	gaat	ct g	attg	agca	g gc	acccc	tga	4	80
a	atgg	gagt	t gc	ugad.	ayac	666	aac+	ttc	tgac	- ggaa	tg q	cgcc	gtgt	g ag	taagg	ĊCC	5	40
C	cgtg	gccg	a ga	aget	ycay	cyc	9400		- , -								6	.00

cggaggccct tttctttgtg caatttgaga agggagagag ctacttccac atgcacgtgc

togtggaaac caccggggtg aaatccatgg ttttgggacg tttcctgagt cagattcgcg

600

660

aaaaactgat tcagagaatt taccgcggga tcgagccgac tttgccaaac tggttcgcgg 720 tcacaaagac cagaaatggc gccggaggcg ggaacaaggt ggtggatgag tgctacatcc 780 840 ccaattactt gctccccaaa acccagcctg agctccagtg ggcgtggact aatatggaac 900 agtatttaag cgcctgtttg aatctcacgg agcgtaaacg gttggtggcg cagcatctga cgcacgtgtc gcagacgcag gagcagaaca aagagaatca gaatcccaat tctgatgcgc 960 cggtgatcag atcaaaaact tcagccaggt acatggagct ggtcgggtgg ctcgtggaca 1020 aggggattac ctcggagaag cagtggatcc aggaggacca ggcctcatac atctccttca 1080 atgcggcctc caactcgcgg tcccaaatca aggctgcctt ggacaatgcg ggaaagatta 1140 tgagcctgac taaaaccgcc cccgactacc tggtgggcca gcagcccgtg gaggacattt 1200 ccagcaatcg gatttataaa attttggaac taaacgggta cgatccccaa tatgcggctt 1260 ccgtctttct gggatgggcc acgaaaaagt tcggcaagag gaacaccatc tggctgtttg 1320 ggcctgcaac taccgggaag accaacatcg cggaggccat agcccacact gtgcccttct 1380 acgggtgcgt aaactggacc aatgagaact ttcccttcaa cgactgtgtc gacaagatgg 1440 tgatctggtg ggaggagggg aagatgaccg ccaaggtcgt ggagtcggcc aaagccattc 1500 tcggaggaag caaggtgcgc gtggaccaga aatgcaagtc ctcggcccag atagacccga 1560 ctcccgtgat cgtcacctcc aacaccaaca tgtgcgccgt gattgacggg aactcaacga 1620 ccttcgaaca ccagcagccg ttgcaagacc ggatgttcaa atttgaactc acccgccgtc 1680 tggatcatga ctttgggaag gtcaccaagc aggaagtcaa agactttttc cggtgggcaa 1740 aggatcacgt ggttgaggtg gagcatgaat tctacgtcaa aaagggtgga gccaagaaaa 1800 gacccgcccc cagtgacgca gatataagtg agcccaaacg ggtgcgcgag tcagttgcgc 1860 1920 agccatcgac gtcagacgcg gaagcttcga tcaactacgc agacaggtac caaaacaaat gttctcgtca cgtgggcatg aatctgatgc tgtttccctg cagacaatgc gagagaatga 1980 atcagaattc aaatatctgc ttcactcacg gacagaaaga ctgtttagag tgctttcccg 2040 tgtcagaatc tcaaccggtt tctgtcgtca aaaaggcgta tcagaaactg tgctacattc 2100 atcatatcat gggaaaggtg ccagacgctt gcactgcctg cgatctggtc aatgtggatt 2160 tggatgactg catctttgaa caataaatga tttaaatcag gtatggctgc cgatggttat 2220 cttccagatt ggctcgagga cactctctct gaaggaataa gacagtggtg gaagctcaaa 2280 2340 cctggcccac caccaccaaa gcccgcagag cggcataagg acgacagcag gggtcttgtg cttcctgggt acaagtacct cggacccttc aacggactcg acaagggaga gccggtcaac 2400 gaggcagacg ccgcggccct cgagcacgac aaagcctacg accggcagct cgacagcgga 2460 gacaacccgt acctcaagta caaccacgcc gacgcggagt ttcaggagcg ccttaaagaa 2520 gatacgtctt ttgggggcaa cctcggacga gcagtcttcc aggcgaaaaa gagggttctt 2580 gaacctctcg gcctggttga ggaacctgtt aagacggctc cgggaaaaaa gaggccggta 2640 gagcactete etgtggagee agacteetee tegggaaceg gaaageeggg eeageageet 2700 gcaagaaaaa gattgaattt tggtcagact ggagacgcag actcagtacc tgaccccag 2760 2820 cctctcggac agccaccagc agccccctct ggtctgggaa ctaatacgat ggctacaggc agtggcgcac caatggcaga caataacgag ggcgccgacg gagtgggtaa ttcctccgga 2880 aattggcatt gcgattccac atggatgggc gacagagtca tcaccaccag cacccgaacc 2940 tgggccctgc ccacctacaa caaccacctc tacaaacaaa tttccagcca atcaggagcc 3000 tcgaacgaca atcactactt tggctacagc accccttggg ggtattttga cttcaacaga 3060 ttccactgcc acttttcacc acgtgactgg caaagactca tcaacaacaa ctggggattc 3120 cgacccaaga gactcaactt caacctcttt aacattcaag tcaaagaggt cacgcagaat 3180 gacggtacga cgacgattgc caataacctt accagcacgg ttcaggtgtt tactgactcg 3240 gagtaccage tecegtacgt ceteggeteg gegeateaag gatgeeteec geegtteeca 3300 gcagacgtct tcatggtgcc acagtatgga tacctcaccc tgaacaacgg gagtcaggca 3360 gtaggacgct cttcatttta ctgcctggag tactttcctt ctcagatgct gcgtaccgga 3420 aacaacttta cetteageta caettttgag gaegtteett tecacageag etaegeteae 3480 agccagagtc tggaccgtct catgaatcct ctcatcgacc agtacctgta ttacttgagc 3540 agaacaaaca ctccaagtgg aaccaccacg cagtcaaggc ttcagttttc tcaggcccca 3600 gccagtgaca ttcgggacca gtctaggaac tggcttcctg gaccctgtta ccgccagcag 3660 cgagtatgaa agacatctgc ggataacaac aacagtgaat actcgtggac tggagctacc 3720 aagtaccacc tcaatggcag agactctctg gtgaatccgg ggcccgccat ggcaagccac 3780 3840 aaggacgatg aagaaaagtt ttttcctcag agcggggttc tcatctttgg gaagcaaggc tcagagaaaa caaatgtgaa cattgaaaag gtcatgatta cagacgaaga ggaaatccca 3900 acaaccaatc ccgtggctac ggagcagtat ggttctgtat ctaccaacct ccagagaggc 3960 aacagacaag cagctaccgc agatgtcaac acacaaggcg ttcttccagg catggtctgg 4020 caggacagag atgtgtacct tcaggggccc atctgggcaa agattccaca cacggacgga 4080 cattttcacc cctctcccct catgggtgga ttcggactta aacaccctcc tccacagatt 4140 4200 ctcatcaaga acaccccggt acctgcgaat ccttcgacca ccttcagtgc ggcaaagttt gcttccttca tcacacagta ctccacggga cacggtcagc gtggagatcg agtgggagct 4260

acadaacdaa	aacagcaaac	gctggaatcc	cgaaattcag	tacacttcca	actacaacaa	4320
-tatattaat	cgtggacttt	accgtggata	ctaatggcgt	gtattcagag	cctcgcccca	4380
geolyceae	atacctgact	cataatctqt	aattgcttgt	taatcaataa	accgtttaat	4440
ttggcaccay	tgaactttgg	tetetacata	tttctttctt	atctagtttc	catggctacg	4500
tcgtttcagt	gcatggcggg	ttootcotta	actacaagga	acccctagtg	atggagttgg	4560
tagataagta	gcatggcggg	Cladicacia	atasaaccaa	gcgaccaaag	gtcgcccgac	4620
ccactccctc	tetgegeget	cgetegetea	Ccgaggoogg	acacaaaaaa	gtcgcccgac	4680
gcccgggctt	tgccccggcg	gcctcagtga	gcgagcgagc	gegeagagag	ggagtgggca	4681
a						

<210> 19 <211> 4683 <212> DNA <213> AAV-6

ttggccactc cctctctgcg cgctcgctcg ctcactgagg ccgggcgacc aaaggtcgcc 60 cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc gagcgcgcag agagggagtg 120 gccaactcca tcactagggg ttcctggagg ggtggagtcg tgacgtgaat tacgtcatag 180 ggttagggag gtcctgtatt agaggtcacg tgagtgtttt gcgacatttt gcgacaccat. 240 gtggtcacgc tgggtattta agcccgagtg agcacgcagg gtctccattt tgaagcggga 300 ggtttgaacg cgcagcgcca tgccggggtt ttacgagatt gtgattaagg tccccagcga 360 ccttgacgag catctgcccg gcatttctga cagctttgtg aactgggtgg ccgagaagga 420 atgggagttg ccgccagatt ctgacatgga tctgaatctg attgagcagg cacccctgac 480 cgtggccgag aagctgcagc gcgacttcct ggtccactgg cgccgcgtga gtaaggcccc 540 ggaggccctc ttctttgttc agttcgagaa gggcgagtcc tacttccacc tccatattct 600 ggtggagacc acgggggtca aatccatggt gctgggccgc ttcctgagtc agattagcga 660 caagctggtg cagaccatct accgcgggat cgagccgacc ctgcccaact ggttcgcggt 720 gaccaagacg cgtaatggcg ccggaggggg gaacaaggtg gtggacgagt gctacatccc 780 caactacctc ctgcccaaga ctcagcccga gctgcagtgg gcgtggacta acatggagga 840 900 gtatataagc gcgtgtttaa acctggccga gcgcaaacgg ctcgtggcgc acgacctgac ccacgtcage cagacccagg agcagaacaa ggagaatctg aaccccaatt ctgacgcgcc 960 tgtcatccgg tcaaaaacct ccgcacgcta catggagctg gtcgggtggc tggtggaccg 1020 gggcatcacc teegagaage agtggateea ggaggaeeag geetegtaea teteetteaa 1080 cgccgcctcc aactcgcggt cccagatcaa ggccgctctg gacaatgccg gcaagatcat 1140 ggcgctgacc aaatccgcgc ccgactacct ggtaggcccc gctccgcccg ccgacattaa 1200 aaccaaccgc atttaccgca teetggaget gaacggetac gaccetgeet acgeeggete 1260 cgtctttctc ggctgggccc agaaaaggtt cggaaaacgc aacaccatct ggctgtttgg 1320 geeggeeace aegggeaaga ecaacatege ggaageeate geecaegeeg tgeeetteta 1380 cggctgcgtc aactggacca atgagaactt tcccttcaac gattgcgtcg acaagatggt 1440 gatctggtgg gaggagggca agatgacggc caaggtcgtg gagtccgcca aggccattct 1500 cggcggcagc aaggtgcgcg tggaccaaaa gtgcaagtcg tccgcccaga tcgatcccac 1560 ccccgtgatc gtcacctcca acaccaacat gtgcgccgtg attgacggga acagcaccac 1620 cttcgagcac cagcagccgt tgcaggaccg gatgttcaaa tttgaactca cccgccgtct 1680 ggagcatgac tttggcaagg tgacaaagca ggaagtcaaa gagttcttcc gctgggcgca 1740 ggatcacgtg accgaggtgg cgcatgagtt ctacgtcaga aagggtggag ccaacaacag 1800 accegecece gatgaegegg ataaaagega geecaagegg geetgeeest eagtegegga 1860 tccatcgacg tcagacgcgg aaggagctcc ggtggacttt gccgacaggt accaaaacaa 1920 atgttctcgt cacgcgggca tgcttcagat gctgtttccc tgcaaaacat gcgagagaat 1980 gaatcagaat ttcaacattt gcttcacgca cgggaccaga gactgttcag aatgtttccc 2040 cggcgtgtca gaatctcaac cggtcgtcag aaagaggacg tatcggaaac tctgtgccat 2100 tcatcatctg ctggggcggg ctcccgagat tgcttgctcg gcctgcgatc tggtcaacgt 2160 ggatctggat gactgtgttt ctgagcaata aatgacttaa accaggtatg gctgccgatg 2220 gttatcttcc agattggctc gaggacaacc tctctgaggg cattcggcag tggtgggact 2280 tgaaacctgg agccccgaaa cccaaagcca accagcaaaa gcaggacgac ggccggggtc 2340 tggtgcttcc tggctacaag tacctcggac ccttcaacgg actcgacaag ggggagcccg 2400 tcaacgcggc ggatgcagcg gccctcgagc acgacaaggc ctacgaccag cagctcaaag 2460 cgggtgacaa teegtacetg eggtataace aegeegaege egagttteag gagegtetge 2520 aagaagatac gtcttttggg ggcaacctcg ggcgagcagt cttccaggcc aagaagaggg 2580 ttctcgaacc ttttggtctg gttgaggaag gtgctaagac ggctcctgga aagaaacgtc 2640 2700 cggtagagca gtcgccacaa gagccagact cctcctcggg cattggcaag acaggccagc agcccgctaa aaagagactc aattttggtc agactggcga ctcagagtca gtccccgacc 2760 cacaacetet eggagaacet ecageaacee eegetgetgt gggacetaet acaatggett 2820 caggcggtgg cgcaccaatg gcagacaata acgaaggcgc cgacggagtg ggtaatgcct 2880 caggaaattg gcattgcgat tccacatggc tgggcgacag agtcatcacc accagcaccc 2940

	gaacatgggc cttgcccacc tataacaacc acctctacaa gcaaatctcc agtgcttcaa	3000
	cgggggccag caacgacaac cactacttcg gctacagcac cccctggggg tattttgatt	3060
	egggggcag caacgacaac cactacticg getacaget 111 3331	3120
·	tcaacagatt ccactgccat ttctcaccac gtgactggca gcgactcatc aacaacaatt	3180
	ggggattccg gcccaagaga ctcaacttca agctcttcaa catccaagtc aaggaggtca	3240
	cgacgaatga tggcgtcacg accatcgcta ataaccttac cagcacggtt caagtcttgt	3300
	cggactcgga gtaccagttc ccgtacgtcc tcggctctgc gcaccagggc tgcctccctc	3360
	cgttcccggc ggacgtgttc atgattccgc agtacggcta cctaacgctc aacaatggca	
	gccaggcagt gggacgctca tccttttact gcctggaata tttcccatcg cagatgctga	3420
	gaacgggcaa taactttacc ttcagctaca ccttcgagga cgtgcctttc cacagcagct	3480
	acgcgcacag ccagagcctg gaccggctga tgaatcctct catcgaccag tacctgtatt	3540
	acctgaacag aactcacaat cagtccggaa gtgcccaaaa caaggacttg ctgtttagcc	3600
	gtgggtctcc agctggcatg tctgttcagc ccaaaaactg gctacctgga ccctgttacc	3660
	ggcagcagcg cgtttctaaa acaaaaacag acaacaacaa cagcaacttt acctggactg	3720
•	gtgcttcaaa atataacctt aatgggcgtg aatctataat caaccctggc actgctatgg	3780
	cctcacacaa agacgacaaa gacaagttct ttcccatgag cggtgtcatg atttttggaa	3840
	aggagagcgc cggagcttca aacactgcat tggacaatgt catgatcaca gacgaagagg	3900
	aggagagege eggageeted dans g	3960
	agagcagcag cacagaccct gcgaccggag atgtgcatgt tatgggagcc ttacctggaa	4020
	agagcagcag cacagaccct gcgaccggag agggtcctat ttgggccaaa attcctcaca	4080
	tggtgtggca agacagagac gtatacctgc agggtcctat ttgggccaaa attcctcaca	4140
	cggatggaca ctttcacccg tctcctctca tgggcggctt tggacttaag cacccgcctc	4200
	ctcagatect cateaaaaac acgeetgtte etgegaatee teeggeagag tttteggeta	4260
	caaagtttgc ttcattcatc acccagtatt ccacaggaca agtgagcgtg gagattgaat	4320
	gggagctgca gaaagaaaac agcaaacgct ggaatcccga agtgcagtat acatctaact	
7. ·	atgcaaaatc tgccaacgtt gatttcactg tggacaacaa tggactttat actgagcctc	4380
	gccccattgg cacccgttac ctcacccgtc ccctgtaatt gtgtgttaat caataaaccg	4440
	gttaattegt gteagttgaa etttggtete atgteettat tatettatet	4500
	gcaaccggtt acacattaac tgcttagttg cgcttcgcga atacccctag tgatggagtt	4560
	geceacteee tetatgegeg etegeteget eggtggggee ggeagageag agetetgeeg	4620
	totgoggaco titggtoogo aggococaco gagogagoga gogogoatag agggagtggo	4680
		4683
	caa	
• •		

```
<210> 20
<211> 16
<212> DNA
<213> rep binding motif
<400> 20
gctcgctcgc tcgctg
```